PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-256003
(43)Date of publication of application : 21.09.2001

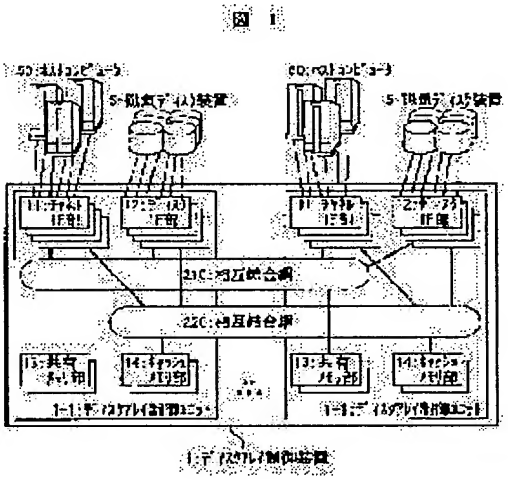
(51)Int.Cl. G06F 3/06
G06F 12/08
G06F 12/16

(21)Application number : 2000-072469
(22)Date of filing : 10.03.2000

(71)Applicant : HITACHI LTD
(72)Inventor : FUJIMOTO KAZUHISA
KANAI HIROKI
FUJIBAYASHI AKIRA
SAKURAI WATARU

(54) DISK ARRAY CONTROLLER, ITS DISK ARRAY CONTROL UNIT AND ITS EXPANDING METHOD

(57)Abstract:
PROBLEM TO BE SOLVED: To solve a problem that lowering of performance in data transition among plural disk array controllers is inevitable when the plural disk array controllers are attempted to be operated as one disk array controller.
SOLUTION: This disk array controller provided with a channel IF part, a disk IF part, a cache memory part and a shared memory part and plural disk array control units to read/write data, has a mutual coupling network to connect the shared memories in the plural disk array control units and a mutual coupling network to connect the cache memory parts in the plural disk array control units are provided.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision

of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

* NOTICES *

1. This document has been translated by computer.So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]
[Claim 1] 1 or two or more channel-interface sections which have an interface with a host computer. An interface with a disk unit. It is the disk array control unit equipped with the above. Ranging over two or more aforementioned disk array control units, it connects through the 1st cross coupling network between the connections of the aforementioned channel-interface section in each aforementioned disk array control unit and the aforementioned disk interface section, and the aforementioned cache memory section. It is characterized by connecting ranging over two or more aforementioned disk array control units through the 2nd cross coupling network between the connections of the aforementioned channel interface in each aforementioned disk array control unit and the aforementioned disk interface, and the aforementioned shared memory section.
[Claim 2] It is the disk array control unit according to claim 1 characterized by connecting mutually the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section in each aforementioned disk array control unit through the cross coupling network of the above 1st, and connecting mutually the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section through the cross coupling network of the above 2nd.
[Claim 3] It is the disk array control unit according to claim 1 which the cross coupling network of the above 1st connects between the aforementioned channel-interface section in each aforementioned disk array control unit and the aforementioned disk interface section, and the aforementioned cache memory section, and is characterized by carrying out the direct file of between the aforementioned channel interchange FE section and the aforementioned disk interface section, and the aforementioned shared memory section within each aforementioned disk array control unit.
[Claim 4] It is the disk array control unit according to claim 1 which it connects through the 3rd interconnection network in a self-unit within each aforementioned disk array control unit between the aforementioned channel-interface section in each aforementioned disk array control unit and the aforementioned disk interface section, and the aforementioned cache memory section, and is characterized by carrying out the direct file of between each aforementioned channel-interface section and two or more aforementioned disk interface sections, and two or more aforementioned shared memory sections within the aforementioned disk array control unit.
[Claim 5] 1 or two or more channel-interface sections which have an interface with a host computer. An interface with a disk unit. It is the disk array control unit equipped with the above, and is characterized by having the 1st connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section, and having the 2nd connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section.
[Claim 6] It is the disk array control unit according to claim 5 which the aforementioned disk array control unit is stored in one case, and is characterized by equipping the above 1st and the

2nd connection path with a connector.
[Claim 7] The disk array control unit characterized by providing the following. 1 or two or more channel-interface sections which have an interface with a host computer. 1 or two or more disk interface sections which have an interface with a disk unit. The cache memory section which stores temporarily the data by which read/write is carried out to the aforementioned disk unit. It has the shared memory section which stores the control information about the data transfer between the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section, and the management information of the aforementioned disk unit. The aforementioned channel-interface section receives the read/write demand of the data from the aforementioned host computer. Data transfer between an interface with the aforementioned host computer and the aforementioned cache memory section is performed. the aforementioned disk interface section By performing data transfer between the aforementioned disk unit and the aforementioned cache memory section It is the disk array control unit which performs read/write of data. It has the 1st connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section. The disk array control unit which has the 2nd connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section two or more And the switching and balancing box out of which each 1st [the] of the aforementioned disk array control unit and the 2nd connection path are connected, combines mutually two or more aforementioned disk array control units, and it cheats as one disk array control unit.
[Claim 8] 1 or two or more channel-interface sections which have an interface with a host computer. 1 or two or more disk interface sections which have an interface with a disk unit. The cache memory section which stores temporarily the data by which read/write is carried out to the aforementioned disk unit. It has the shared memory section which stores the control information about the data transfer between the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section, and the management information of the aforementioned disk unit. The aforementioned channel-interface section receives the read/write demand of the data from the aforementioned host computer. Data transfer between an interface with the aforementioned host computer and the aforementioned cache memory section is performed. the aforementioned disk interface section By performing data transfer between the aforementioned disk unit and the aforementioned cache memory section It is the disk array control unit which performs read/write of data. It has the 1st connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section. The disk array control unit which has the 2nd connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section, And the switching and balancing box out of which the 1st of the aforementioned disk array control unit and the 2nd connection path are connected, combines mutually two or more aforementioned disk array control units, and it cheats as one disk array control unit. It has a connector in the above 1st, the 2nd connection path, and the aforementioned switching and balancing box, it is the extension method of the disk array control unit equipped with the above, the aforementioned switching and balancing box is constituted so that the connector of the 1st of two or more aforementioned disk array control units and the 2nd connection path may be connected, and it carries out extending the aforementioned disk array control unit as the feature by connecting the connector of the above 1st and the 2nd connection path, and the connector of the aforementioned switching and balancing box with a cable.
[Claim 9] It is the extension method of the disk array control unit according to claim 8 which carries out [that the aforementioned switching and balancing box adds the effective aforementioned correspondence information about the aforementioned disk array control unit

extended by the aforementioned table in connection with having the table which stores the correspondence information which shows correspondence with the information which specifies the aforementioned disk array control unit which should be connected with the address of the destination of the data which pass along the aforementioned switching and balancing box, and extending the aforementioned disk array control unit, and] as the feature.
[Claim 10] The aforementioned disk array control unit is the extension method of the disk array control unit of the claim 8 characterized by connecting the case which was stored in the case, contained the aforementioned switching and balancing box and the aforementioned disk array control unit to a certain case, and contained other disk array control units by connecting two or more cables to the aforementioned switching and balancing box, respectively.
[Claim 11] 1 or two or more channel-interface sections which have an interface with a host computer. An interface with a disk unit. Are the disk array control unit equipped with the above, and it has the 1st connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section. Two or more disk array control units which have the 2nd connection path for connection with other aforementioned disk array control units in the connection of the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section, And it is characterized by storing the switching and balancing box out of which the 1st of the aforementioned disk array control unit and the 2nd connection path are connected, combines mutually two or more aforementioned disk array control units, and it cheats as one disk array control unit in one case.

[Translation done.]

* NOTICES *

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[The technical field to which invention belongs] this invention relates to the control unit of the disk array equipment which stores data in two or more magnetic disk units.
[0002]
[Description of the Prior Art] Compared with the I/O performance of the primary storage of the computer which uses a semiconductor memory as a storage, the disk-subsystem (henceforth "subsystem") I/O performance which uses a magnetic disk as a storage is small about 3-4 figures, and efforts to raise contracting this difference from the former, i.e., the I/O performance of a subsystem, are made. As one method for raising the I/O performance of a subsystem, a subsystem is constituted from two or more magnetic disk units, and the system which stores data in two or more magnetic disk units and which is called so-called disk array is known.
[0003] For example, two or more channel IF sections 11 which perform data transfer between a host computer 50 and the disk array control unit 2 with the conventional technology as shown in drawing 2 , Two or more disk IF sections 12 which perform data transfer between a magnetic disk unit 5 and the disk array control unit 2, The cache memory section 14 which stores the data of a magnetic disk unit 5 temporarily, the control information (for example, the information about the data transfer control between the channel IF section 11 and the disk IF section 12, and the cache memory section 14 —) about the disk array control unit 2 Have the shared memory section 13 which stores the management information of the data stored in a magnetic disk unit 5, and it sets in one disk array control unit 2. The shared memory section 13 and the cache memory section 14 have accessible composition from all the channel IF sections 11 and disk IF sections 12. In this disk array control unit 2, it connects with the cross coupling network 21 and the cross coupling network 22, respectively between the channel IF section 11 and the disk IF section 12, and the shared memory section 13 and between the channel IF section 11 and the disk IF section 12, and the cache memory section 14.
[0004] The channel IF section 11 has the microprocessor (not shown) which controls the interface for connecting with a host computer 50, and the I/O over a host computer 50. Moreover, the disk IF section 12 has the microprocessor (not shown) which controls the interface for connecting with a magnetic disk unit 5, and the I/O over a magnetic disk unit 5. Moreover, the disk IF section 12 also performs execution of a RAID function.
[0005] In this conventional disk array control unit 2, when there was an upper limit in the capacity of the disk in which connection with one disk array control unit 2 is possible and one disk array control unit 2 needed to record the data more than the ** support ****** record amount of data, two or more disk array control units 2 were installed, and the channel was connected to two or more disk array control units 2 from the host computer 50.
[0006] Moreover, when the host computer 50 more than the number of host channels connectable with one disk array control unit 2 needed to be connected, two or more disk array control units 2 were installed, and the host computer 50 was connected to each.
[0007] Moreover, when making data shift among two or more disk array control units 2, the channel was connected to both two disk array control units 2 which shift data from one host computer 50, and data were shifted through the host computer 50.
[0008] moreover, with U.S. JP,5,B and the conventional technology currently indicated by No. 680 or 640 As shown in drawing 3 , when shifting data between two disk array control units 3, between two disk array control units 3 Connect some interfaces (drawing 2) with a host computer 50 with the data shift path 8, and the data shift path 8 is minded for the data of the magnetic disk unit 5 connected with one disk array control unit 3. It had shifted to the magnetic disk unit 5 connected with another disk array control unit 3.
[0009] Moreover, when the data more than the record amount of data which one disk array control unit supports need to be recorded with other conventional technology, When the host computer more than the number of host channels connectable with one disk array control unit again needs to be connected, Moreover, when making data shift among two or more disk array control units, As shown in drawing 4 , two or more disk array control units 4 were installed, and the interface with those host computers 50 was connected to the host computer 50 through the cross coupling network 23 which consists of switches.
[0010] Two director equipments share a shared memory in JP,11-66693,A, and the array disk processor restored when the data spindle which constitutes a disk array becomes out of control is shown in it. Here, composition which installs two or more array disks is not shown.
[0011]
[Problem(s) to be Solved by the Invention] It is in the inclination which cuts down the costs which employment of a computer system and a storage system, maintenance, and management take by centralizing the computer and storage which were being conventionally distributed to every place in a data center, and constituting a computer system and a storage system from a big business represented by a bank, a security, the telephone company, etc.
[0012] The support (KONEKUTIBITI) of the channel interface for connecting with hundreds of or more sets of host computers and the support with a storage capacity of hundreds of terabytes or more are demanded of the disk array control unit large-sized/high-end in such an inclination.
[0013] On the other hand, the demand to the disk array control unit of small-scale composition (small case) of having had the same high efficiency and high-reliability as a large-sized/high-end disk array control unit is increasing by expansion of open market in recent years, and the spread of the storage area networks (SAN) expected from now on.
[0014] To the former demand, how to connect two or more conventional large-sized/high-end disk array control units, and constitute an overly large-scale disk array control unit can be considered.
[0015] Moreover, to the latter demand, how to constitute the equipment which set to the model (for example, thing which lessened the number of the channel IF section and the disk IF sections) of the minimum configuration of the conventional large-sized/high-end disk array control unit, and miniaturized the case to it can be considered. Moreover, how to constitute the equipment which supports the composition of the middle-scale shell large-scale which the conventional disk array control unit is supporting can be considered by connecting two or more sets of this miniaturized equipment.
[0016] a disk array control unit -- high efficiency and quantity same from small-scale composition to overly large-scale composition as mentioned above -- the control unit of composition of that there is a scalability which can respond by reliable architecture is needed, and, for that, the disk array control unit which are two or more disk array control units, and can be employed as one disk array control unit is needed
[0017] However, with the conventional technology shown in drawing 2 , although it was possible to have increased the number of channels connected to a host computer 50 by increasing simply the number of the disk array control unit 2 and storage capacity, when one host computer 50 needed to store data in two or more disk array control units 2, the host computer 50 needed to connect the channel to all the disk array control units 2. Moreover, when data were accessed, the data to access need to grasp whether it is stored in the magnetic disk unit 5 connected to which disk array control unit 2, and needed to access the host computer 50 by specifying the target disk array control unit 2.
[0018] Therefore, it was difficult to employ two or more sets of disk array control units as one

disk array control unit.
[0019] it was possible to have accessed the data stored in the magnetic disk unit 5 to which three between disk array control units is connected by the data shift path 8, and which accumulates, only connects a host computer 50 to one set of the disk array control unit 3, and is connected with other disk array control units 3 with the conventional technology shown in drawing 3 , and it was possible to have employed two or more disk array control units 3 as one disk array control unit
[0020] However, from a host computer 50, there is a lead demand of data, and when there are no data in the magnetic disk unit 5 connected to the disk array control unit 3, the data shift path 8 is disk array minded [3]. From the disk array control unit 3 by which the magnetic disk unit 5 which sends the lead demand and stores the corresponding data was connected to other disk array control units 3, demand data needed to be received through the data shift path 8, and demand data needed to be returned to the host computer 50. Therefore, when a certain host computer 50 accessed the data stored in the magnetic disk unit 5 connected with disk array control units 3 other than disk array control unit 3 with which confidence is connected, there was a problem that a performance fell remarkably.
[0021] Moreover, when making the data stored in the magnetic disk unit connected to a different disk array control unit from the disk array control unit with which this host computer is connected shift to the magnetic disk unit 5 connected to the disk array control unit 3 with which this host computer 50 is connected beforehand, in order to make data shift through the data shift path 8 in data with the frequency high in order to prevent the above-mentioned problem accessed from a certain host computer 50, there was a problem that where of a performance is low.
[0022] Moreover, it is possible for a host computer 50 to access all the disk array control units 4 through the cross coupling network 23 which used the switch with the conventional technology shown in drawing 4 .
[0023] however, in order to employ two or more disk array control units 4 as one disk array control unit The data of all the disk array control units 4 connected to the switch into the switch which constitutes the cross coupling network 23 When it is necessary to have the map in which it is shown in which disk array control unit 4 it is stored and there is an access demand from a host computer 50, a command is analyzed in a switch and the function assigned to the disk array control unit 4 which stores demand data is needed.
[0024] In this case, since it is necessary to analyze a command also in the switch connected on it in addition to the command analysis in the conventional channel IF section 11, there is a problem that a performance falls, compared with the case where the direct file of the host computer 50 is carried out to the disk array control unit 4.
[0025] There were the following functions in a high-end disk array control unit. Namely, the duplicate of a certain business-use data set (it corresponds to a logical volume) is held. On the usual business, data are simultaneously updated to an original data set and a duplicate data set. For example, when there is a demand which takes backup of the data set, Renewal of data is stopped about a duplicate data set, it is used for a backup rise, and in an original data set, when business is continued and backup is completed, there is a function to take the adjustment of an original data set and the data set of a duplicate.
[0026] In the conventional technology shown in drawing 2 - drawing 4 , when it is going to hold the duplicate of a data set between disk array control units which are different when realizing the above-mentioned function, in order to make a data set shift between disk array control units, there was a problem that a performance fell remarkably.
[0027] The purpose of this invention is to offer the disk array control unit of composition of that there is a scalability which can respond by the architecture of the high efficiency and high-reliability same from small-scale composition to overly large-scale composition.
[0028] More specifically, the purpose of this invention is to stop degradation and realize offering the disk array system which can stop the degradation by the data shift between two or more disk array control units, and the function which a disk array control unit has by two or more sets of disk array control units, when it is going to employ two or more sets of disk array control units

as one disk array control unit.
[0029]
[Means for Solving the Problem] 1 or two or more channel-interface sections in which the above-mentioned purpose has an interface with a host computer, 1 or two or more disk interface sections which have an interface with a disk unit, The cache memory section which stores temporarily the data by which read/write is carried out to the aforementioned disk unit, It has the shared memory section which stores the control information about the data transfer between the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section, and the management information of the aforementioned disk unit. The aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section are connected. The aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section are connected. As opposed to the read/write demand of the data from the aforementioned host computer the aforementioned channel-interface section Data transfer between an interface with the aforementioned host computer and the aforementioned cache memory section is performed. the aforementioned disk interface section By performing data transfer between the aforementioned disk unit and the aforementioned cache memory section The disk array control unit equipped with the function as a disk controller in itself performs read/write of data A means to be the disk array control unit which ********** and to connect between the aforementioned shared memory sections in two or more aforementioned disk array control units, It has a means to connect between the aforementioned cache memory sections in two or more aforementioned disk array control units. From the aforementioned channel-interface section and the aforementioned disk interface section in the aforementioned disk array control unit It is attained by the disk array control unit characterized by read/write being possible in the data of the aforementioned shared memory section in other aforementioned disk array control units, or the data of the aforementioned cache memory section.
[0030]
[Embodiments of the Invention] Hereafter, although explained taking the case of a magnetic disk unit as mass data storage equipment, it may not be restricted to a magnetic disk as large capacity storage, and you may be large capacity storage like DVD.
[0031] As one of the gestalten of operation of this invention, preferably Between two or more aforementioned channel-interface sections in two or more aforementioned disk array control units and two or more aforementioned disk interface sections, and two or more aforementioned cache memory sections It connects with the cross coupling network using the switch over between two or more aforementioned disk array control units. It connects with the cross coupling network using the switch over between two or more aforementioned disk array control units between two or more aforementioned channel-interface sections and two or more aforementioned disk interface sections, and two or more aforementioned shared memory sections.
[0032] Moreover, between two or more aforementioned channel-interface sections in two or more desirable aforementioned disk array control units and two or more aforementioned disk interface sections, and two or more aforementioned cache memory sections It connects with the cross coupling network using the switch over between two or more aforementioned disk array control units. Between two or more aforementioned channel-interface sections and two or more aforementioned disk interface sections, and two or more aforementioned shared memory sections Within the aforementioned disk array control unit, a direct file is carried out and between the aforementioned shared memory sections is connected between the aforementioned disk array control units with the cross coupling network using the switch over between two or more aforementioned disk array control units.
[0033] Moreover, between two or more aforementioned channel-interface sections in two or more desirable aforementioned disk array control units and two or more aforementioned disk interface sections, and two or more aforementioned cache memory sections Within the aforementioned disk array control unit, it connects with the interconnection network which used

the switch in a self-unit. among two or more aforementioned disk array control units Between the aforementioned cache memory sections is connected with the cross coupling network using the switch over between two or more aforementioned disk array control units. Between two or more aforementioned channel-interface sections and two or more aforementioned disk interface sections, and two or more aforementioned shared memory sections Within the aforementioned disk array control unit, a direct file is carried out and between the aforementioned shared memory sections is connected among two or more aforementioned disk array control units with the cross coupling network using the switch over between two or more aforementioned disk array control units.

[0034] Moreover, it is as follows if its attention is paid to the read/write of the data of a host computer and a magnetic disk unit. The channel-interface section which has an interface with a host computer, The disk interface section which has an interface with a magnetic disk unit, The cache memory section which stores temporarily the data by which read/write is carried out to the aforementioned magnetic disk unit, The shared memory section which stores the control information about the data transfer between the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section, and the management information of the aforementioned magnetic disk unit, A means to connect the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned cache memory section, It has a means to connect the aforementioned channel-interface section and the aforementioned disk interface section, and the aforementioned shared memory section. As opposed to the read/write demand of the data from the aforementioned host computer the aforementioned channel-interface section Data transfer between an interface with the aforementioned host computer and the aforementioned cache memory section is performed. the aforementioned disk interface section By performing data transfer between the aforementioned magnetic disk unit and the aforementioned cache memory section A means to be the disk array control unit which ********** the disk array control unit which performs read/write of data, and to connect between the aforementioned shared memory sections in two or more aforementioned disk array control units, Have a means to connect between the aforementioned cache memory sections in two or more aforementioned disk array control units, and these connecting means are minded. Read/write is possible in the data of the aforementioned magnetic disk unit connected only to the different aforementioned disk array control unit from this disk array control unit from the aforementioned host computer connected only to the one aforementioned disk array control unit.

[0035] In addition, the technical problem which this application indicates, and its solution method are clarified with the column and drawing of an operation gestalt of invention.

[0036] Hereafter, the example of this invention is explained using a drawing.

[0037] One example of this invention is shown in [example 1] drawing 1 , drawing 7 , and drawing 8 . In the following examples, although the cross coupling network has explained the thing using the switch to the example, it connects mutually, and control information and data should just be transmitted, for example, it may consist of buses.

[0038] As shown in drawing 1 , the disk array control unit 1 consists of two or more disk array control units 1-1. The disk array control unit 1-1 The interface section 11 with a host computer 50 (channel IF section), The interface section 12 with a magnetic disk unit 5 (disk IF section), Have the shared memory section 13 and the cache memory section 14, and it connects through the cross coupling network 210 over two or more disk array control units 1-1 between the channel IF section 11 and the disk IF section 12, and the shared memory section 13. It connects through the cross coupling network 220 over two or more disk array control units 1-1 between the channel IF section 11 and the disk IF section 12, and the cache memory section 14. That is, it has accessible composition from all the channel IF sections 11 and disk IF sections 12 through the cross coupling network 210 or the cross coupling network 220 to all the shared memory sections 13 or all the cache memory sections 14.

[0039] Although one disk array control unit may be constituted as one case or it may be constituted as a module, it has the function as one disk controller in itself. It explains as what constitutes the respectively separate case from drawing 7 .

[0040] A concrete example in one disk array control unit 1-1 is shown in drawing 7 . The disk array control unit 1-1 Two channel IF sections 11 with a host computer 50, Two disk IF sections 12 with a magnetic disk unit 5, and two switches 110 for a shared memory path (SM-SW), Two path switches (CM-SW) 111 for cache memories, and the two shared memory sections 13, It has the two cache memory sections 14, shared memory (SM) access paths 135 and 136, cache memory (CM) access paths 137 and 138, SM paths 141 between cases, and the CM path 142 between cases. it has come out that SM path between cases says SM path between disk array control units and CM path between cases as CM path between disk array units

[0041] The channel IF section 11 Two IF102 with a host computer 50 (host IF), Two microprocessors 101 which control the I/O over a host computer 50, The access-control section 104 which controls access to the shared memory section 13 (SM access-control section), It has the access-control section (CM access-control section) 105 which controls access to the cache memory section 14, and the data transfer between a host computer 50 and the cache memory section 14 and a transfer of the control information between a microprocessor 101 and the shared memory section 13 are performed. A microprocessor 101 and a host IF 102 are connected by internal bus 106, and the direct file of the CM access-control section 105 is carried out to two hosts IF 102. Moreover, the direct file of the SM access-control section 104 is carried out to two microprocessors 101.

[0042] The disk IF section 12 Two IF103 with a magnetic disk unit 5 (drive IF), Two microprocessors 101 which control the I/O over a magnetic disk unit 5, The one access-control section 104 to the shared memory section 13 (SM access-control section), It has the one access-control section (CM access-control section) 105 to the cache memory section 14, and the data transfer between a magnetic disk unit 5 and the cache memory section 14 and a transfer of the control information between a microprocessor 101 and the shared memory section 13 are performed. A microprocessor 101 and drive IF 103 are connected by internal bus 106, and the direct file of the CM access-control section 105 is carried out to two drives IF 103. Moreover, the direct file of the SM access-control section 104 is carried out to two microprocessors 101. The disk IF section also performs execution of a RAID function.

[0043] The shared memory section 13 has the shared memory (SM) controller 107 and a memory module 109, and stores the control information (for example, information about the data transfer control between the channel IF section 11 and the disk IF section 12, and the cache memory section 14, management information of the data recorded on a magnetic disk unit 5) of the disk array control unit 1-1 etc.

[0044] The cache memory section 14 has the cache memory (CM) controller 108 and a memory module 109, and stores temporarily the data recorded on a magnetic disk unit 5.

[0045] Two SM access paths 135 are connected to SM access-control section 104, and they are connected to two different SM-SW110, respectively. Two access paths 136 are connected to SM-SW110, and they are connected to two different SM controllers 107, respectively. Therefore, every one of a total of two access paths 136 is connected to the SM controller 107 from two SM-SW110. By carrying out like this, the access root from one SM access-control section 104 to one SM controller 107 is set to two. Since it becomes possible to access to the shared memory section 13 by another access root by this even when an obstacle occurs in one access path or SM-SW110, obstacle-proof nature can be raised.

[0046] Two CM access paths 137 are connected to CM access-control section 105, and they are connected to two different CM-SW111, respectively. Two access paths 138 are connected to CM-SW111, and they are connected to two different CM controllers 108, respectively. Therefore, every one of a total of two access paths 138 is connected to the CM controller 108 from two CM-SW111. By carrying out like this, the access root from one CM access-control section 105 to one CM controller 108 is set to two. Since it becomes possible to access to the cache memory section 14 by another access root by this even when an obstacle occurs in one access path or CM-SW111, obstacle-proof nature can be raised.

[0047] Every one of a total of four SM access paths 135 is connected to SM-SW110, respectively from two channel IF sections 11 and two disk IF sections 12. Moreover, a total of two of every one access path 136 to the two shared memory sections 13 are connected to SM-

SW110. Moreover, the-two SM path 141 between cases for connecting with SM-SW110 of other disk array control units 1-1 is connected. One side may be [an input and another side] the objects for an output, and this performs a transfer of the information on both directions [each].

[0048] Every one of a total of four CM access paths 137 is connected to CM-SW111, respectively from two channel IF sections 11 and two disk IF sections 12. Moreover, a total of two of every one access path 138 to the two cache memory sections 14 are connected to CM-SW111. Moreover, the-two CM path 142 between cases for connecting with CM-SW111 of other disk array control units 1-1 is connected. These two have the same property as SM path between cases.
[0049] In SM-SW110 or CM-SW111 Since the above access paths are connected, within SM-SW110 or CM-SW111 The demand from four access paths from the channel IF section 11 and the disk IF section 12 Two access paths to the shared memory section 13 or the cache memory section 14 in the self-disk array control unit 1-1, It has the function distributed to the access path between [of two] cases to the shared memory section 13 or the cache memory section 14 in the other disk array control unit 1-1.
[0050] SM-SW110 is the connection of the channel IF section 11 and the disk IF section 12, and the shared memory section 13 in drawing 7 , and CM-SW111 is the connection of the channel IF section 11 and the disk IF section 12, and the cache memory section 14.
[0051] The example of the disk array control unit 1 which connected two disk array control units 1-1 shown in drawing 7 is shown in drawing 8 .
[0052] Each SM path 141 between cases is connected through SM-SW121 between cases between each disk array control unit 1-1 SM-SW110 in -1 and 1-1-2.
[0053] Similarly, each CM path 142 between cases is connected through CM-SW122 between cases between each disk array control unit 1-1 CM-SW111 in -1 and 1-1-2. SW 121 and 122 is mounted as a switching and balancing box.
[0054] It is satisfactory, when the number of the disk array control units 1-1 to connect is two and this invention is carried out like this example, even if it connects the path between direct cases, without minding between [SW / 121 and 122] cases. By doing so, it becomes possible to cut down the overhead of the data transfer processing generated between [SW / 121 and 122] cases, and a performance improves.
[0055] It becomes possible to connect three or more sets of the disk array control units 1-1 by increasing from the example which shows the number of paths connectable between [SW / 121 and 122] cases to drawing 8 . Since there is a physical limitation in the number of paths which can be mounted between [SW / 121 and 122] cases, when the disk array control unit 1-1 to connect is increased, it becomes impossible moreover, to connect among [SW] one case. In such a case, it is connecting between [SW] cases to multi-stage, and it becomes possible [increasing the number of the connectable disk array control unit 1-1]. The example of mounting which carried out cross coupling of three sets of the disk array control units by the switching and balancing box is shown in drawing 19 . It mentions later for details.
[0056] Moreover, when connecting three or more sets of the disk array control units 1-1, as shown in drawing 13 , between SM-SW110 in each disk array control unit 1-1 and CM-SW111 is connected in the shape of a loop. With such composition, it becomes possible to connect two or more disk array control units 1-1, without using between [SW / 121 and 122] cases.
[0057] In this case, what is necessary is just to connect SM paths between cases which have come out of each disk array control unit 1-1, and CM paths between cases by the connector. Although omitted drawing, if a connector is prepared in the portion besides the disk array control unit 1-1 of the SM path 141 between cases, and the CM path 142 between cases, it is convenient for extending a disk array control unit. SM-SW110, the SM path 141 between cases, and SM-SW121 between cases constitute the cross coupling network 210 of drawing 1 from drawing 8 , and the cross coupling network 220 of drawing 1 consists of CM-SW111, a CM path 142 between cases, and CM-SW122 between cases.
[0058] In drawing 8 , an example in the case of reading data from the host computer 50 connected to the disk array control unit 1-1-1 is described.

[0059] First, a host computer 50 publishes the read-out demand of data in the channel IF section 11 in disk array control unit 1-1-1 to which self is connected. The microprocessor 101 in the channel IF section 11 which received the demand investigates in which magnetic disk unit 5 both two disk array control units 1-1-1 and the shared memory section 13 in 1-1-2 are accessed, and the demanded data are stored. It can investigate in which magnetic disk unit 5 the translation table to which it is made to correspond in which magnetic disk unit 5 the address and data of demand data are stored is stored in the shared memory section 13, and the demanded data are stored.
[0060] Next, the microprocessor 101 in the channel IF section 11 which received the demand accesses both two disk array control units 1-1-1 and the shared memory section 13 in 1-1-2, and checks whether the data required of each disk array control unit 1-1-1 and the cache memory section 14 in 1-1-2 are stored. The directory information on the data in the cache memory section 14 is stored in the shared memory section 13, and it can check whether demand data exist in each cache memory section 14.
[0061] When data are in the disk array control unit 1-1 cache memory section 14 of -1 by that cause, the data is transmitted to the channel IF section 11 through own CM-SW111, and is sent to a host computer 50. When data are in the disk array control unit 1-1 cache memory section 14 of -2, the data is transmitted to the channel IF section 11 through CM-SW111 in disk array control unit 1-1-2, CM-SW122 between cases, and own CM-SW111, and is sent to a host computer 50.
[0062] When data exist in [no] the cache memory section 14, to the microprocessor 101 in the disk IF section 12 to which the magnetic disk unit 5 in which demand data are stored is connected, a microprocessor 101 reads demand data, and it publishes an instruction so that it may store in the cache memory section 14. The microprocessor 101 in the disk IF section 12 which received the instruction reads data from the magnetic disk unit 5 in which demand data are stored, and stores demand data in one of the cache memory sections 14 two disk array control units 1-1-1 and of 1-1-2. When it stores data in the cache memory section 14 in disk array control unit 1-1-2 with which the magnetic disk unit 5 in which demand data are stored is connected By minding CM-SW111 in the disk array control unit 1-1-2 Moreover, when it stores data in the cache memory section 14 in disk array control unit 1-1-1 which is different in the disk array control unit 1-1-2 with which the magnetic disk unit 5 in which demand data are stored is connected By minding each CM-SW111 and CM-SW122 between cases, data are transmitted to the cache memory section 14.
[0063] The microprocessor 101 in the disk IF section 12 tells the cache memory section 14 which stored data in the microprocessor 101 in the channel IF section 11 which published the instruction, after storing demand data in the cache memory section 14. The microprocessor 101 in the channel IF section 11 which received it reads data from the cache memory section 14 in which data were stored, and sends them to a host computer 50.
[0064] Without being conscious of whether the host computer 50 is stored in the magnetic disk unit 5 to which demand data are connected with which disk array control unit 1-1 according to this example, an access demand is only published to the disk array control unit 1-1 with which self is connected, it becomes possible to perform the writing and read-out of data, and it becomes possible to a host computer 50 to show two or more sets of the disk array control units 1-1 as one disk array control unit 1.
[0065] Moreover, since data can be read through an internal cross coupling network and the internal cache memory section 14 when reading data from the magnetic disk unit 5 connected with a disk array control unit 1-1 which is different in the disk control unit 1-1 which received the demand, the need of shifting data through the channel IF section 11 of both disk array control units 1-1 is lost, and it becomes possible to stop read-out of data, and the performance degradation of writing.
[0066] Other examples of this invention are shown in [example 2] drawing 5 , drawing 9 , and drawing 10 .
[0067] As shown in drawing 5 , the composition of the disk array control unit 1 which consists of two or more disk array control units 1-2 is the same as the composition shown in drawing 1 of

an example 1 except for the connection composition between the channel IF section 11 and the disk IF section 12, and the shared memory section 13.
[0068] Between the channel IF section 11 and the disk IF section 12, and the shared memory section 13, it connects directly into the disk array control unit 1-2. Moreover, among two or more disk array control units 1-2, the shared memory section 13 is connected through the cross coupling network 24.
[0069] As mentioned above, in this example, it becomes possible by carrying out the direct file of the channel IF section 11 and the disk IF section 12, and the shared memory section 13 into the disk array control unit 1-2 to shorten the access time to the shared memory section 13 compared with the case where it connects through the cross coupling network 210 shown in the example 1.
[0070] A concrete example in one disk array control unit 1-2 is shown in drawing 9 .
[0071] It is the same as that of the composition which also shows the composition in the disk array control unit 1-2 to drawing 7 of an example 1 except for the connection composition between the channel IF section 11 and the disk IF section 12, and the shared memory section 13.
[0072] a disk array -- a control unit -- one -- - -- two -- a host computer -- 50 -- two -- a ** -- a channel -- IF -- the section -- 11 -- a magnetic disk unit -- five -- two -- a ** -- a disk -- IF -- the section -- 12 -- two -- a ** -- a cache memory -- ** -- a path -- a switch (CM-SW) -- 111 -- two -- a ** -- a shared memory -- the section -- 13 -- two -- a ** -- a cache memory -- the section -- 14 -- a
[0073] Two SM access paths 139 are connected to SM access-control section 104, and they are connected to two different SM controllers 107, respectively. Therefore, every one of a total of four SM access paths 139 is connected to the SM controller 107 from two channel IF sections 11 and two disk IF sections 12. Moreover, the-two SM path 143 between cases for connecting with the SM controller 107 of other disk array control units 1-2 is connected.
[0074] By the SM controller 107, since the above access paths are connected, within the SM controller 107, it has the function to distribute the demand from four SM access paths 139 from the channel IF section 11 and the disk IF section 12 to the access path to a memory module 109, and the SM access path 143 between [of two] cases to the shared memory section 13 in the other disk array control unit 1-2.
[0075] The example of the disk array control unit 1 which connected two disk array control units 1-2 shown in drawing 9 is shown in drawing 10 .
[0076] Between the shared memory sections 13 in each disk array control unit 1-2, each SM path 143 between cases is connected through SM-SW121 between cases.
[0077] Except it, it is the same as that of the composition shown in drawing 8 of an example 1. In this case, the SM controller 107 makes the connection of the channel IF section, the disk IF section, and the shared memory section.
[0078] It is satisfactory, when the number of the disk array control units 1-2 to connect is two and this invention is carried out like an example 1, even if it connects the path between direct cases, without minding between [SW / 121 and 122] cases. By doing so, it becomes possible to cut down the overhead of the data transfer processing generated between [SW / 121 and 122] cases, and a performance improves.
[0079] Moreover, it becomes possible to connect three or more sets of the disk array control units 1-2 by increasing from the example which shows the number of paths connectable between [SW / 121 and 122] cases to drawing 10 like an example 1. Since there is a physical limitation in the number of paths which can be mounted between [SW / 121 and 122] cases, when the disk array control unit 1-2 to connect is increased, it becomes impossible moreover, to connect among [SW] one case. In such a case, it is connecting between [SW] cases to multi-stage, and it becomes possible [increasing the number of the connectable disk array control unit 1-2]. This is mounted as a switching and balancing box like an example 1.
[0080] Moreover, when connecting three or more sets of the disk array control units 1-2, in an example 1, it becomes possible to connect two or more disk array control units 1-2, without using between [SW / 121 and 122] cases by taking the composition shown in drawing 13 , and

the composition of the shape of same loop.
[0081] In this example, operation of each part within the disk array control unit 1-2 in the case of performing read-out/writing of the data from a host computer 50 to a magnetic disk unit 5 is the same as that of an example 1 except for access being performed to the shared memory section 13 in other disk array control units 1-2 through the shared memory section 13 and the cross coupling network 24 in the self-disk array control unit 1-2. In addition, the cross coupling network 24 consists of an SM path 143 between cases, SM-SW121 between cases, and an SM controller 107.
[0082] Without being conscious of whether the host computer 50 is stored in the magnetic disk unit 5 to which demand data are connected with which disk array control unit 1-2 according to this example, an access demand is only published to the disk array control unit 1-2 with which self is connected, it becomes possible to perform the writing and read-out of data, and it becomes possible to a host computer 50 to show two or more sets of the disk array control units 1-2 as one disk array control unit 1.
[0083] Moreover, since data can be read through an internal cross coupling network and the internal cache memory section 14 when reading data from the magnetic disk unit 5 connected with a disk array control unit 1-2 which is different in the disk control unit 1-2 which received the demand, the need of shifting data through the channel IF section 11 of both disk array control units 1-2 is lost, and it becomes possible to stop read-out of data, and the performance degradation of writing.
[0084] Other examples of this invention are shown in [example 3] drawing 6 , drawing 11 , and drawing 12 .
[0085] As shown in drawing 6 , the composition of the disk array control unit 1 which consists of two or more disk array control units 1-3 is the same as the composition shown in drawing 5 of an example 2 except for the connection composition between the channel IF section 11 and the disk IF section 12, and the cache memory section 14.
[0086] Between the channel IF section 11 and the disk IF section 12, and the cache memory section 14, it connects through the cross coupling network 22 into the disk array control unit 1-3. Moreover, among two or more disk array control units 1-3, the cache memory section 14 is connected through the cross coupling network 25. It is the following reason to connect through the cross coupling network 22 to the direct file of the channel IF section 11 and the disk IF section 12, and the shared memory section 13 being carried out like an example 2 between the channel IF section 11 and the disk IF section 12, and the cache memory section 14. By the cache memory section 14, data are treated in several K bytes of unit to the control information treated in the shared memory section 13 being several bytes. Therefore, raising a throughput is planned, connecting with the number of pins restricted through the cross coupling network 22.
[0087] As mentioned above, when making data shift between the cache memory sections 14 of a different disk array control unit 1-3 by separating the cross coupling network 22 which connects the channel IF section 11 and the disk IF section 12, and the cache memory section 14, and the cross coupling network 24 which connects the cache memory section 14 between the disk array control units 1-3, barring access to the cache memory section 14 of the access demand from a host computer 50 is lost. The disk IF section manages shift of the data between cache memories. The function to which data are made to shift between the cache memory sections 14 of a different disk array control unit 1-3 is a function required in order to move data to a disk array control unit with low access frequency, when access from a host computer 50 concentrates to one disk array control unit 1-3.
[0088] A concrete example in one disk array control unit 1-3 is shown in drawing 11 .
[0089] It is the same as that of the composition which also shows the composition in the disk array control unit 1-3 to drawing 9 of an example 2 except for the connection composition between the channel IF section 11 and the disk IF section 12, and the cache memory section 14. Here, the CM controller 108 makes the connection of the channel IF section, the disk IF section, and the cache memory section.
[0090] a disk array -- a control unit -- one -- - -- three -- a host computer -- 50 -- two -- a ** -- a channel -- IF -- the section -- 11 -- a magnetic disk unit -- five -- two -- a ** -- a

disk — IF — the section — 12 — two — a ** — a cache memory — ** — a path — a switch (CM-SW) — 111 — two — a ** — a shared memory — the section — 13 — two — a ** — a cache memory — the section — 14 — a

[0091] Two CM access paths 137 are connected to CM access-control section 105, and they are connected to two different CM-SW111, respectively. Two access paths 138 are connected to CM-SW111, and they are connected to two different CM controllers 108, respectively. Therefore, every one of a total of two access paths 138 is connected to the CM controller 108 from two CM-SW111. Moreover, the-two CM path 144 between cases for connecting with the CM controller 108 of other disk array control units 1-3 is connected.

[0092] By the CM controller 108, since the above access paths are connected, within the CM controller 108, it has the function to distribute the demand from two CM access paths 138 from CM-SW111 to the access path to a memory module 109, and the CM access path 144 between [of two] cases to the cache memory section 14 in the other disk array control unit 1-3.

[0093] Every one of a total of four CM access paths 137 is connected to CM-SW111, respectively from two channel IF sections 11 and two disk IF sections 12. Moreover, a total of two of every one access path 138 to the two cache memory sections 14 are connected to CM-SW111.

[0094] In CM-SW111, since the above access paths are connected, within CM-SW111, it has the function to distribute the demand from four CM access paths 137 from the channel IF section 11 and the disk IF section 12 to two CM access paths 138 to the cache memory section 14.

[0095] The example of the disk array control unit 1 which connected two disk array control units 1-3 shown in drawing 11 is shown in drawing 12 .

[0096] Between the cache memory sections 14 in each disk array control unit 1-3, each CM path 144 between cases is connected through CM-SW122 between cases. Except it, it is the same as that of the composition shown in drawing 10 of an example 2.

[0097] It is satisfactory, when the number of the disk array control units 1-3 to connect is two and this invention is carried out like an example 2, even if it connects the path between direct cases, without minding between [SW / 121 and 122] cases. By doing so, it becomes possible to cut down the overhead of the data transfer processing generated between [SW / 121 and 122] cases, and a performance improves.

[0098] Moreover, it becomes possible to connect three or more sets of the disk array control units 1-3 by increasing from the example which shows the number of paths connectable between [SW] cases to drawing 12 like an example 2. Since there is a physical limitation in the number of paths which can be mounted between [SW / 121 and 122] cases, when the disk array control unit 1-3 to connect is increased, it becomes impossible moreover, to connect among [SW] one case. In such a case, it is connecting between [SW] cases to multi-stage, and it becomes possible [increasing the number of the connectable disk array control units 121 and 122].

[0099] Moreover, when connecting three or more sets of the disk array control units 1-3, in an example 1, it becomes possible to connect two or more disk array control units 1-3, without using between [SW / 121 and 122] cases by taking the composition shown in drawing 13 , and the composition of the shape of same loop.

[0100] In this example, operation of each part within the disk array control unit 1-3 in the case of performing read-out/writing of the data from a host computer 50 to a magnetic disk unit 5 is the same as that of an example 2 except for access being performed to the cache memory section 14 in other disk array control units 1-3 through the cache memory section 14 and the cross coupling network 25 in the self-disk array control unit 1-3.

[0101] Without being conscious of whether the host computer 50 is stored in the magnetic disk unit 5 to which demand data are connected with which disk array control unit 1-3 according to this example, an access demand is only published to the disk array control unit 1-3 with which self is connected, it becomes possible to perform the writing and read-out of data, and it becomes possible to a host computer 50 to show two or more sets of the disk array control units 1-3 as one disk array control unit 1.

[0102] Moreover, since data can be read through an internal cross coupling network and the

internal cache memory section 14 when reading data from the magnetic disk unit 5 connected with a disk array control unit 1-3 which is different in the disk control unit 1-3 which received the demand, the need of shifting data through the channel IF section 11 of both disk array control units 1-3 is lost, and it becomes possible to stop read-out of data, and the performance degradation of writing.

[0103] Next, the example of use of the example of this invention is explained. There are the following functions in a high-end disk array control unit. Namely, the duplicate of a certain business-use data set (it corresponds to a logical volume) is held. On the usual business, data are simultaneously updated to an original data set and a duplicate data set. For example, when there is a demand which takes backup of the data set, Renewal of data is stopped about a duplicate data set, it is used for a backup rise, and an original data set has the function to take the adjustment of the data of an original data set and the data set of a duplicate, when business is continued and backup is completed.

[0104] In the disk array control unit 1 shown in an example 1, when holding the duplicate of a data set between different disk array control units 1-1, how to realize the above-mentioned function is explained using drawing 8 .

[0105] Here, suppose that an original data set is stored in the magnetic disk unit 5 temporarily connected with the disk array control unit 1-1-1 shown in drawing 8 , and the data set of a duplicate is stored in the magnetic disk unit 5 connected with the disk array control unit 1-1-2. Moreover, the host computer 50 connected with the disk array control unit 1-1-1 will perform the usual business, and the work which takes backup of data will be done on the tape unit (not shown) to which the host computer 50 connected with the disk array control unit 1-1-2 is connected with self.

[0106] On the usual business, when there is a write request of data to the target data set from the host computer 50 connected with the disk array control unit 1-1-1, the microprocessor 101 in the disk array control unit 1-1 channel IF section 11 connected with the host computer 50 connected with -1 transmits and writes the data sent from the host computer 50 in the cache memory section 14 of the disk array control unit 1-1-1. The aforementioned microprocessor 101 to next, the microprocessor 101 in the disk IF section 12 with which the magnetic disk unit 5 in which the original data set is stored is connected An instruction is published through the shared memory section 13 of the disk array control unit 1-1-1. Data are read from the cache memory section 14 of the disk array control unit 1-1-1, and it transmits to the disk IF section 12 connected with the magnetic disk unit 5 in which the original data set is stored, and it transmits to a magnetic disk unit 5, and is made to write in it from there.

[0107] The microprocessor 101 in the disk array control unit 1-1 channel IF section 11 of -1 is supervising the renewal of data of an original data set, and stores the control information which shows the amount of renewal of data of an original data set in the disk array control unit 1-1 shared memory section 13 of -1. If it becomes beyond the value as which the amount of renewal of data was determined beforehand, the aforementioned microprocessor 101 will publish an instruction to the microprocessor 101 in the disk IF section 12 with which the magnetic disk unit 5 in which the original data set is stored is connected so that the content of updating of an original data set may be reflected in the data set of a duplicate. In response to it, a microprocessor 101 reads the data updated from the magnetic disk unit 5, and changes the address of the updating data into the address of the data set of a duplicate. The updating data is transmitted and written in the cache memory section 14 of the disk array control unit 1-1-2 through CM-SW111 of the disk array control unit 1-1-1, CM-SW122 between cases, and CM-SW111 of the disk array control unit 1-1-2. Next, updating data are read from the cache memory section 14, and it transmits to the disk IF section 12 connected with the magnetic disk unit 5 in which the data set of a duplicate is stored, and from there, it transmits to a magnetic disk unit 5, and writes in it.

[0108] By the above-mentioned operation, an original data set and the data set of a duplicate are usually held in business.

[0109] When there is a demand which takes backup to the target data set from the host computer 50 connected with -2, the microprocessor 101 in the disk array control unit 1-1

channel IF section 11 with which the host computer 50 is connected publishes an instruction so that the renewal of data to the data set of a duplicate may be interrupted through the shared memory section 13 of -1 temporarily to the microprocessor 101 in the disk array control unit 1-1 channel IF section 11 with which the host computer 50 which is usually performing business is connected. The microprocessor 101 which received it interrupts renewal of data temporarily. Next, the microprocessor 101 in the channel IF section 11 connected with the host computer 50 which advanced the demand of backup To the microprocessor 101 in the disk IF section 12 connected with the magnetic disk unit 5 in which the data set of a duplicate is stored An instruction is published through the shared memory section 13 of the disk array control unit 1-1-2. The data set of a duplicate is read from a magnetic disk unit 5, and it transmits to the disk IF section 12, and it transmits to the cache memory section 14 of the disk array control unit 1-1-2, and is made to write in it from the disk IF section 12. After it is completed, the microprocessor 101 in the aforementioned channel IF section 11 reads the data set of a duplicate from the cache memory section 14 of the disk array control unit 1-1-2, transmits it to the channel IF section 11, and is sent to the host computer 50 which advanced the demand of backup there.

[0110] After backup of a data set is completed, the microprocessor 101 in the channel IF section 11 connected with the host computer 50 which advanced the demand of backup publishes an instruction through the shared memory section 13 of the disk array control unit 1-1-1 to the microprocessor 101 in the channel IF section 11 with which the host computer 50 which usually performs business is connected, and makes the data updated during the backup processing in an original data set reflect in the data set of a duplicate. This method is the same as the method reflecting the updating data in the above-mentioned usual business. Since according to this example data can be shifted through an internal cross coupling network and the internal cache memory section 14 between two disk array control units 1-1-1 and 1-1-2 when realizing the above-mentioned function, the need of shifting data through both disk array control units 1-1-1 and the channel IF section of 1-1-2 is lost, and it becomes possible to stop degradation when performing the above-mentioned function. Therefore, reducing the efficiency of a user's usual business is lost.

[0111] Also in the disk array control unit 1 of the composition of an example 2 and an example 3, when carrying out this example, it is satisfactory, and the same effect as this example is acquired.

[0112] There are the following as other examples of use. It is necessary to read data from the disk array control unit by which the magnetic disk unit 5 which stores the corresponding data through a cross coupling network when there are no data in the magnetic disk unit 5 which there is a lead demand of data and was connected to the disk array control unit was connected to the disk array control unit from the host computer 50, and to return demand data to a host computer 50 in the disk array control unit 1 shown in an example 1, an example 2, and an example 3. Thus, when performing read/write of data by straddling between disk array control units and making data shift, a performance falls compared with the case where that is not right.

[0113] The function which shifts to the magnetic disk unit 5 connected to the disk array control unit with which this host computer 50 is connected in the data stored in the magnetic disk unit 5 connected to a different disk array control unit from the disk array control unit with which this host computer 50 is connected in data with the frequency high in order to suppress the above data shift accessed from a certain host computer 50 is needed.

[0114] In the disk array control unit 1 shown in an example 1, the method of the above-mentioned data shift is explained using drawing 8 .

[0115] The microprocessor 101 in the channel IF section 11 is supervising the access frequency to the data set in all the magnetic disk units 5 (it corresponds to a logical volume), and stores the control information which shows the access frequency of the aforementioned data set in the shared memory section 13 in the disk array control unit 1-1-1 [same] as self.

[0116] Here from the host computer 50 connected with the disk array control unit 1-1-1 Access concentrates on the data set in the disk array control unit 1-1 magnetic disk unit 5 connected with -2. When access frequency exceeds the value defined beforehand, the microprocessor 101

in the disk array control unit 1-1 channel IF section 11 of -1 To the microprocessor 101 in the disk IF section 12 with which the magnetic disk unit 5 which stores the corresponding data set is connected An instruction is published through the shared memory section 13 in disk array control unit 1-1-2, and it transmits to the cache memory section 14 of the disk array control unit 1-1-2, and is made to read the corresponding data set and to write in it.

[0117] Next, the microprocessor 101 in the disk array control unit 1-1 channel IF section 11 of -1 reads the data which correspond from the cache memory section 14 of the disk array control unit 1-1-2, and transmits them to the cache memory section 14 of the disk array control unit 1-1-1. Next, an instruction is published to the microprocessor 101 in the disk array control unit 1-1 disk IF section 12 of -1 through the shared memory section 13 in disk array control unit 1-1-1, the data which correspond from the cache memory section 14 of the disk array control unit 1-1-1 are read to it, and a magnetic disk unit 5 is made to write in it. Since according to this example data shift can be performed between two disk array control units 1-1 through an internal cross coupling network and the internal cache memory section 14 when performing the above data shift, the need of shifting data through the channel IF section of both disk array control units is lost, and it becomes possible to stop degradation when performing the above-mentioned data shift. Therefore, reducing the efficiency of a user's usual business is lost.

[0118] Also in the disk array control unit of the composition of an example 2 and an example 3, when carrying out this example, it is satisfactory, and the same effect as this example is acquired.

[0119] Next, the example of mounting of examples 1, 2, and 3 is described. Drawing 14 shows an example which carried the disk array control unit 1-1 shown in drawing 7 of an example 1 in the case 201.

[0120] The channel IF section 11 shown in drawing 7 is mounted on the channel IF package (PK) 311, the disk IF section 12 is mounted on a disk IFPK312, SM-SW110 and CM-SW111 are mounted on a switch PK320, and the shared memory section 13 and the cache memory section 14 are mounted on memory PK330. Moreover, on the back plane 340, the SM access paths 135 and 136 and the CM access paths 137 and 138 are wired, and it has become the form which puts each above PK in a back plane 340.

[0121] The cable for SM path 141 between cases and the cable for CM path 142 between cases are connected to a switch PK320, and the other end of each cable is connected to the connectors 221 and 222 in the side of a case 201, respectively. Illustration of a cable is omitted. 350 is a wall box and supplies power to above PK. Thus, the disk array control unit is equipped with the function of a disk array control unit as one case.

[0122] Drawing 15 shows an example in the case of connecting two cases 201 shown in drawing 14 .

[0123] SM-SW121 between cases shown in drawing 8 and CM-SW122 between cases are carried in the switching and balancing box 210. The CM path 142 between cases with which the SM path 141 between cases which leads to SM-SW121 between cases leads to a connector 221 at CM-SW122 between cases is connected to the connector 222.

[0124] When connecting two sets of cases 201 and 202, it is a cable 231 about the connector 221 for SM path between cases of a case 201, and the connector 221 of a switching and balancing box 210, and a cable 232 ties the connector 222 for CM path between cases of a case 201, and the connector 222 of a switching and balancing box 210. Similarly, a cable 231 ties the connector 221 of a case 202, and the connector 221 of a switching and balancing box 210, and a cable 232 ties the connector 222 of a case 202, and the connector 222 of a switching and balancing box 210.

[0125] By things, it becomes possible to support the number of connection channels to the host computer which cannot be supported by one case, or the storage capacity which cannot be supported by one case making it be the above.

[0126] Here, a case 201 can be made into a basic case, a case 202 can be made into the case for extension, and a switching and balancing box 210 can also be carried in a case 202. The installation space of a switching and balancing box 210 can be deleted without raising the manufacturing cost of the basic case 201 by carrying out like this.

[0127] Applying this example to the disk array control unit of an example 2 or an example 3 does not have any problem.
[0128] The topology in the case of connecting three sets of disk array control units to drawing 18 by SM-SW121 between cases and CM-SW122 between cases is shown. It solves and what has a large capacity is needed from each of this case where two sets of disk array control units are connected since the number of connection of SW is increasing as mentioned above. Like illustration, the SM path 141 between cases and the CM path 142 between cases connect with SM-SW121 between cases, and CM-SW122 between cases, respectively, and each disk array control unit 1-1-1 to 1-1-3 functions as one disk array control unit as a whole with them.
[0129] The example of mounting is shown in drawing 19 . Here, the switching and balancing box 210 is mounted as another case. A case 201,202,203 is connected to this by a connector 221 and the connector 222 through the cable 231 for SM path between cases, and the cable 232 for CM path between cases, respectively. If the capacity and the connector which connect four or more disk array control units to a switching and balancing box 210 are prepared, extension from after is easy.
[0130] Drawing 21 shows the format of the data which pass along a switching and balancing box 210. Data take the form of a packet and consist of a destination address 401, the command section 402, and data division 403. The address is the address on a shared memory and a cache memory.
[0131] Drawing 22 shows the switch-off substitute table 410 for the switching prepared in the switching and balancing box 210. Correspondence of the number of a disk array control unit including a destination address and its address is memorized here. A switching and balancing box 210 asks for the change place of a switch with reference to this switch-off substitute table from the address 401 of a packet 400, and performs change control of a switch.
[0132] When extending a disk array control unit, it is based on the following procedure. If an excess is in the connector which extends a disk array control unit to a switching and balancing box 210, a cable 231 and a cable 232 will be connected to the connector. When there was no excess, after connecting a switching and balancing box to multi-stage, a cable 231 and a cable 232 are connected to the connector. It adds by the disk array control unit which extended the address of the switch-off substitute table 410 in a switching and balancing box 210, and port No. with it. When the aforementioned address is written in beforehand and extended, there is also a method of standing an effective flag.
[0133] Drawing 20 shows other examples of connection. Three disk array control units are connected in series like illustration. At this time, SM-SW110 and CM-SW111 have the bridge function to transmit the inputted information to other disk array control units as it is. Although it has SM-SW110 and CM-SW111 drawing instead, it is good also as a bus structure. And you may combine two or more disk array control units by the common bus which made bus connection.
[0134] Next, the example of mounting of further others is shown. As shown in drawing 16 , the number of sheets of the package (PK) of the disk array control unit 1-1 carried in the case 201 shown in drawing 14 is reduced, and it considers as the case 205 which carried the disk array control unit 1-1 of a minimum configuration.
[0135] Since middle-scale by carrying two or more cases 205 and switching and balancing boxes 210 in one case 206, and connecting through a switching and balancing box 210 by the method which showed between cases 205 in the example 6, and the same method as shown in drawing 17 , it becomes possible to constitute the disk array control unit of large-scale composition.
[0136] In addition, you may take the form called module that what is necessary is just a thing with the function of the disk array control unit on which it crawled, and which has been described as case 205, without taking the form of a case. Moreover, it is suitably decided as a matter on mounting whether give a wall box for every disk array control unit in the case of drawing 17 or supply electric power from a common wall box.
[0137]
[Effect of the Invention] According to this invention, when it is going to employ two or more sets of disk array control units as one disk array control unit, it becomes possible to offer the disk array system which stops the degradation by the data shift between two or more disk array control units.

[Translation done.]

http://www4.ipdl.jpo.go.jp/cgi-bin/tran_web_cgi_ejje 2003/06/10

* NOTICES *

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1] Drawing showing the composition of the disk array control unit by this invention.
[Drawing 2] Drawing showing the composition of the conventional disk array control unit.
[Drawing 3] Drawing showing other composition of the conventional disk array control unit.
[Drawing 4] Drawing showing other composition of the conventional disk array control unit.
[Drawing 5] Drawing showing other composition of the disk array control unit by this invention.
[Drawing 6] Drawing showing other composition of the disk array control unit by this invention.
[Drawing 7] Drawing showing the detailed composition in the disk array control unit shown in drawing 1 .
[Drawing 8] Drawing showing the composition which connects two or more disk array control units shown in drawing 7 .
[Drawing 9] Drawing showing the detailed composition in the disk array control unit shown in drawing 5 .
[Drawing 10] Drawing showing the composition which connects two or more disk array control units shown in drawing 9 .
[Drawing 11] Drawing showing the detailed composition in the disk array control unit shown in drawing 6 .
[Drawing 12] Drawing showing the composition which connects two or more disk array control units shown in drawing 11 .
[Drawing 13] Drawing showing other composition which connects two or more disk array control units shown in drawing 7 .
[Drawing 14] Drawing showing the example of loading to the case of the disk array control unit by this invention.
[Drawing 15] Drawing showing the composition which connects two or more cases which carried the disk array control unit by this invention.
[Drawing 16] Drawing showing other examples of loading to the case of the disk array control unit by this invention.
[Drawing 17] Drawing showing the example which carries the disk array control unit by this invention to two or more sets and one case.
[Drawing 18] Drawing showing the wiring structure of connecting three sets of disk array control systems with the switch between cases.
[Drawing 19] Drawing showing the example of 1 mounting of the wiring structure of drawing 18 .
[Drawing 20] Drawing showing other examples for connecting three or more disk array control units by this invention.
[Drawing 21] Drawing showing an example of the data format of the information given to a switching and balancing box.
[Drawing 22] Drawing showing an example of the switch-off substitute table prepared in the switching and balancing box.
[Description of Notations]
1: A disk array control unit, 1-1 [— The channel IF section 12 / — The disk IF section 13 / — The shared memory section, 14 / — 210 The cache memory section 220 / — A cross coupling

network, 50 / — Host computer] — A disk array control unit, 5 — A magnetic disk unit, 11

[Translation done.]

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（A）　(11)特許出願公開番号
特開2001－256003
（P2001－256003A）
(43)公開日　平成13年９月21日(2001.9.21)

(51)Int.Cl.⁷	識別記号	FI	テーマコード(参考)
G06F 3/06	302	G06F 3/06	302Z 5B005
	540		540 5B018
12/08		12/08	J 5B065
			G
	320		320

審査請求　未請求　請求項の数11　ＯＬ　（全 33 頁）　最終頁に続く

(21)出願番号　特願2000－72469(P2000－72469)

(22)出願日　平成12年３月10日(2000.3.10)

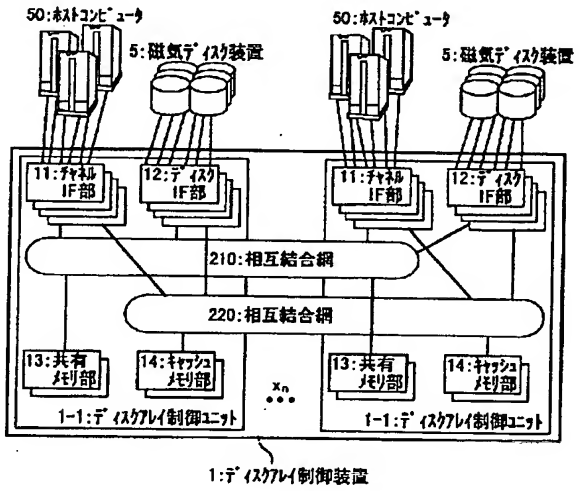
(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(72)発明者　藤本　和久
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内
(72)発明者　金井　宏樹
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内
(74)代理人　100068504
弁理士　小川　勝男　（外１名）

最終頁に続く

(54)【発明の名称】　ディスクアレイ制御装置、そのディスクアレイ制御ユニットおよびその増設方法

(57)【要約】
【課題】複数台のディスクアレイ制御装置を１つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置間でのデータ移行における性能低下が避けられなかった。
【課題を解決するための手段】チャネルＩＦ部と、ディスクＩＦ部と、キャッシュメモリ部と、共有メモリ部とを有し、データのリード／ライトを行うディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置において、複数のディスクアレイ制御ユニット内の共有メモリ部間を接続する相互結合網と、複数のディスクアレイ制御ユニット内のキャッシュメモリ部間を接続する相互結合網を有する。

図 1



【特許請求の範囲】

【請求項１】ホストコンピュータとのインターフェースを有する１または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する１または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置であって、前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部間は第１の相互結合網を介して複数の前記ディスクアレイ制御ユニットにまたがって接続され、前記各ディスクアレイ制御ユニット内の前記チャネルインターフェースおよび前記ディスクインターフェースと前記共有メモリ部との接続部間は第２の相互結合網を介して複数の前記ディスクアレイ制御ユニットにまたがって接続されていることを特徴とするディスクアレイ制御装置。

【請求項２】前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部は互いに前記第１の相互結合網を介して接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部は互いに前記第２の相互結合網を介して接続されることを特徴とする請求項１記載のディスクアレイ制御装置。

【請求項３】前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部と前記のキャッシュメモリ部との間は、前記第１の相互結合網によって接続され、前記チャネルインターフェー部及び前記ディスクインターフェース部と前記共有メモリ部との間は、前記各ディスクアレイ制御ユニット内では直接接続されていることを特徴とする請求項１記載のディスクアレイ制御装置。

【請求項４】前記各ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間は、前記各ディスクアレイ制御ユニット内では、自ユニット内の第３の相互接続網を通して接続され、前記各チャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記ディスクアレイ制御ユニット内では、直接接続されていることを特徴とする請求項１記載のディスクアレイ制御装置。

【請求項５】ホストコンピュータとのインターフェースを有する１または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する１または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第１の接続パスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第２の接続パスを有することを特徴とするディスクアレイ制御ユニット。

【請求項６】前記ディスクアレイ制御ユニットは１つの筐体に収められ、前記第１、第２の接続パスはコネクタを備えたことを特徴とする請求項５記載のディスクアレイ制御ユニット。

【請求項７】ホストコンピュータとのインターフェースを有する１または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する１または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キ

ャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第１の接続パスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第２の接続パスを有するディスクアレイ制御ユニットが複数個、および前記ディスクアレイ制御ユニットのそれぞれの第１、第２の接続パスが接続され、前記複数個のディスクアレイ制御ユニットを互いに結合し一つのディスクアレイ制御装置とせしめるスイッチボックスとを備えたことを特徴とするディスクアレイ制御装置。

【請求項８】ホストコンピュータとのインターフェースを有する１または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する１または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第１の接続パスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第２の接続パスを有するディスクアレイ制御ユニット、および前記ディスクアレイ制御ユニットの第１、第２の接続パスが接続され、前記複数個のディスクアレイ制御ユニットを互いに結合し一つのディスクアレイ制御装置とせしめるスイッチボックスとを備えたディスクアレイ制御装置において、前記第１、第２の接続パスおよび前記スイッチボックスにコネクタを有し、前記スイッチボックスは複数の前記ディスクアレイ制御ユニットの第１、第２の接続パスのコネクタが接続されるよう構成され、ケーブルによって前記第１、第２の接続パスのコネクタと前記スイッチボックスのコネクタを接続することにより前記ディスクアレイ制御ユニットを増設することを特徴とするディスクアレイ制御ユニットの増設方法。

【請求項９】前記スイッチボックスは前記スイッチボックスを通るデータのあて先のアドレスと接続すべき前記ディスクアレイ制御ユニットを特定する情報との対応を示す対応情報を格納するテーブルを持ち、前記ディスクアレイ制御ユニットを増設するのに伴い、前記テーブルに増設された前記ディスクアレイ制御ユニットに関する有効な前記対応情報を追加することを特徴とする請求項８記載のディスクアレイ制御ユニットの増設方法。

【請求項１０】前記ディスクアレイ制御ユニットはそれぞれ筐体に収められ、ある筐体に前記スイッチボックスと前記ディスクアレイ制御ユニットを収納し、前記スイッチボックスに複数のケーブルを接続することにより他のディスクアレイ制御ユニットを収納した筐体を接続することを特徴とする請求項８のディスクアレイ制御ユニットの増設方法。

【請求項１１】ホストコンピュータとのインターフェースを有する１または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する１または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部は前記ホストコンピュータからのデータのリード／ライト要求に対し、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットであって、前記チャネルインターフェース部および前記ディスクインターフェース部と前記キャッシュメモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第１の接続パスを有し、前記チャネルインターフェース部および前記ディスクインターフェース部と前記共有メモリ部との接続部に他の前記ディスクアレイ制御ユニットへの接続のための第２の接続パスを有する複数のディスクアレイ制御ユニット、および前記ディスクアレイ制御ユニットの第１、第２の接続パスが接続され、前記複数個のディスクアレイ制御ユニットを互いに結合し一つのディスクアレイ制御装置とせしめるスイッチボックスとが１つの筐体に収められていることを特徴とするディスクアレイ制御装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、データを複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置に関する。

【０００２】

【従来の技術】半導体記憶装置を記憶媒体とするコンピュータの主記憶のＩ／Ｏ性能に比べて、磁気ディスクを記憶媒体とするディスクサブシステム（以下「サブシス

テム」という。）のＩ／Ｏ性能は３～４桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのＩ／Ｏ性能を向上させる努力がなされている。サブシステムのＩ／Ｏ性能を向上させるための１つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを複数の磁気ディスク装置に格納する、いわゆるディスクアレイと呼ばれるシステムが知られている。

【０００３】例えば、従来技術では、図２に示すようにホストコンピュータ５０とディスクアレイ制御装置２との間のデータ転送を実行する複数のチャネルＩＦ部１１と、磁気ディスク装置５とディスクアレイ制御装置２間のデータ転送を実行する複数のディスクＩＦ部１２と、磁気ディスク装置５のデータを一時的に格納するキャッシュメモリ部１４と、ディスクアレイ制御装置２に関する制御情報（例えば、チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間のデータ転送制御に関する情報、磁気ディスク装置５に格納するデータの管理情報）を格納する共有メモリ部１３とを備え、１つのディスクアレイ制御装置２内において、共有メモリ部１３およびキャッシュメモリ部１４は全てのチャネルＩＦ部１１及びディスクＩＦ部１２からアクセス可能な構成となっている。このディスクアレイ制御装置２では、チャネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１３との間、及びチャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間は、相互結合網２１、及び相互結合網２２でそれぞれ接続される。

【０００４】チャネルＩＦ部１１は、ホストコンピュータ５０と接続するためのインターフェース及びホストコンピュータ５０に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクＩＦ部１２は、磁気ディスク装置５と接続するためのインターフェース及び磁気ディスク装置５に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクＩＦ部１２は、RAID機能の実行も行う。

【０００５】この従来のディスクアレイ制御装置２では、１つのディスクアレイ制御装置２への接続可能なディスクの容量には上限があり、１つのディスクアレイ制御装置２ががサポート出きる記録データ量以上のデータを記録する必要がある場合、ディスクアレイ制御装置２を複数台設置し、ホストコンピュータ５０から複数のディスクアレイ制御装置２にチャネルを接続していた。

【０００６】また、１つのディスクアレイ制御装置２に接続できるホストチャネル数以上のホストコンピュータ５０を接続する必要がある場合、ディスクアレイ制御装置２を複数台設置し、それぞれにホストコンピュータ５０を接続していた。

【０００７】また、複数のディスクアレイ制御装置２間でデータを移行させる場合、１つのホストコンピュータ５０からデータの移行を行う２つのディスクアレイ制御装置２の両方にチャネルを接続し、ホストコンピュータ５０を介して、データの移行を行っていた。

【０００８】また、米国特許５、６８０、６４０号に開示されている従来技術では、図３に示すように２つのディスクアレイ制御装置３間でデータの移行を行う場合、２つのディスクアレイ制御装置３の間で、ホストコンピュータ５０とのインターフェースの一部（図では２本）をデータ移行パス８で接続し、一方のディスクアレイ制御装置３に繋がる磁気ディスク装置５のデータを、データ移行パス８を介して、もう一方のディスクアレイ制御装置３に繋がる磁気ディスク装置５に移行していた。

【０００９】また、他の従来技術では、１つのディスクアレイ制御装置がサポートする記録データ量以上のデータを記録する必要がある場合、 また１つのディスクアレイ制御装置に接続できるホストチャネル数以上のホストコンピュータを接続する必要がある場合、また、複数のディスクアレイ制御装置間でデータを移行させる場合、図４に示すようにディスクアレイ制御装置４を複数台設置し、それらのホストコンピュータ５０とのインターフェースをスイッチから構成される相互結合網２３を介して、ホストコンピュータ５０に接続していた。

【００１０】特開平１１－６６６９３号公報には、２つのディレクタ装置が共有メモリを共用して、ディスクアレイを構成するデータスピンドルが制御不能になったとき復元するアレイディスク処理装置が示されている。ここではアレイディスクを複数台設置するような構成は示されていない。

【００１１】

【発明が解決しようとする課題】銀行、証券、電話会社等に代表される大企業では、従来各所に分散していたコンピュータ及びストレージを、データセンターの中に集中化してコンピュータシステム及びストレージシステムを構成することにより、コンピュータシステム及びストレージシステムの運用、保守、管理に要する費用を削減する傾向にある。

【００１２】このような傾向の中で、大型／ハイエンドのディスクアレイ制御装置には、数百台以上のホストコンピュータへ接続するためのチャネルインターフェースのサポート(コネクティビティ)、数百テラバイト以上の記憶容量のサポートが要求されている。

【００１３】一方、近年のオープン市場の拡大、今後予想されるストレージ・エリア・ネットワーク(SAN)の普及により、大型／ハイエンドのディスクアレイ制御装置と同様の高機能・高信頼性を備えた小規模構成(小型筐体)のディスクアレイ制御装置への要求が高まっている。

【００１４】前者の要求に対しては、従来の大型／ハイエンドのディスクアレイ制御装置を複数接続して超大規模なディスクアレイ制御装置を構成する方法が考えられる。

【００１５】また後者の要求に対しては、従来の大型／ハイエンドのディスクアレイ制御装置の最小構成のモデル、(例えばチャネルＩＦ部、ディスクＩＦ部の数を少なくしたもの)において筐体を小型化した装置を構成する方法が考えられる。また、この小型化した装置を複数台接続することにより、従来のディスクアレイ制御装置がサポートしている中規模から大規模の構成をサポートする装置を構成する方法が考えられる。

【００１６】ディスクアレイ制御装置では、上記のように、小規模な構成から超大規模な構成まで、同一の高機能・高信頼なアーキテクチャで対応可能な、スケーラビリティのある構成の制御装置が必要となってきており、そのためには、複数のディスクアレイ制御装置で、１つのディスクアレイ制御装置として運用できるディスクアレイ制御装置が必要となる。

【００１７】しかし、図２に示す従来技術では、ディスクアレイ制御装置２の台数を単純に増やすことによってホストコンピュータ５０に接続するチャネル数、記憶容量を増やすことが可能であるが、１つのホストコンピュータ５０が複数のディスクアレイ制御装置２にデータを格納する必要がある場合、ホストコンピュータ５０は全てのディスクアレイ制御装置２に対してチャネルを接続する必要があった。また、データにアクセスする場合、ホストコンピュータ５０はアクセスするデータがどのディスクアレイ制御装置２に接続された磁気ディスク装置５に格納されているかを把握し、目的のディスクアレイ制御装置２を特定してアクセスを行う必要があった。

【００１８】したがって、複数台のディスクアレイ制御装置を１つのディスクアレイ制御装置として運用することは困難であった。

【００１９】図３に示す従来技術では、データ移行パス８によりディスクアレイ制御装置間３が接続されているため、ホストコンピュータ５０は１台のディスクアレイ制御装置３に接続するだけで、他のディスクアレイ制御装置３に繋がる磁気ディスク装置５に格納されたデータにアクセスすることが可能であり、複数のディスクアレイ制御装置３を１つのディスクアレイ制御装置として運用することが可能であった。

【００２０】しかし、ホストコンピュータ５０からディスクアレイ制御装置３に、例えばデータのリード要求があり、そのディスクアレイ制御装置３に接続された磁気ディスク装置５にデータが無かった場合、データ移行パス８を介して、他のディスクアレイ制御装置３にそのリード要求を送り、該当するデータを格納している磁気ディスク装置５が接続されたディスクアレイ制御装置３から、データ移行パス８を介して要求データを受け取り、ホストコンピュータ５０へ要求データを返す必要があった。したがって、あるホストコンピュータ５０が、自信が繋がるディスクアレイ制御装置３以外のディスクアレイ制御装置３に繋がる磁気ディスク装置５に格納されたデータにアクセスする場合、著しく性能が低下するという問題があった。

【００２１】また、上記問題を防ぐため、あるホストコンピュータ５０からアクセスされる頻度の高いデータの中で、該ホストコンピュータが繋がるディスクアレイ制御装置とは異なるディスクアレイ制御装置に接続された磁気ディスク装置に格納されているデータを、予め該ホストコンピュータ５０が繋がるディスクアレイ制御装置３に接続された磁気ディスク装置５に移行させておく場合、データ移行パス８を介してデータを移行させるため、性能が低いという問題があった。

【００２２】また、図４に示す従来技術では、スイッチを用いた相互結合網２３を介して、ホストコンピュータ５０は全てのディスクアレイ制御装置４にアクセスすることが可能である。

【００２３】しかし、複数のディスクアレイ制御装置４を１つのディスクアレイ制御装置として運用するためには、相互結合網２３を構成するスイッチ内に、そのスイッチに接続された全てのディスクアレイ制御装置４のデータが、どのディスクアレイ制御装置４に格納されているかを示すマップを持つ必要があり、ホストコンピュータ５０からアクセス要求があった場合、スイッチにおいてコマンドを解析し、要求データを格納しているディスクアレイ制御装置４に割り振る機能が必要となる。

【００２４】この場合、従来のチャネルＩＦ部１１でのコマンド解析に加え、その上に繋がるスイッチにおいてもコマンドを解析する必要があるため、ホストコンピュータ５０がディスクアレイ制御装置４に直接接続されている場合に比べ、性能が低下するという問題がある。

【００２５】ハイエンドのディスクアレイ制御装置には以下のような機能があった。すなわち、ある業務用のデータセット(論理ボリュームに対応)の複製を保持しておき、通常の業務ではオリジナルデータセットと複製データセットに対して同時にデータの更新を行い、例えば、そのデータセットのバックアップをとる要求があった場合、複製データセットについてデータの更新を中止し、それをバックアップアップ用に使用し、オリジナルデータセットでは業務を継続し、バックアップが終了した時点で、オリジナルデータセットと複製のデータセットの整合性をとるという機能がある。

【００２６】図２～図４に示す従来技術において、上記機能を実現する場合、異なるディスクアレイ制御装置間でデータセットの複製を保持しようとした場合には、ディスクアレイ制御装置間でデータセットを移行させる必要があるため、性能が著しく低下するという問題があった。

【００２７】本発明の目的は、小規模な構成から超大規模な構成まで、同一の高機能・高信頼性のアーキテクチャで対応可能な、スケーラビリティのある構成のディスクアレイ制御装置を提供することにある。

【００２８】より具体的には、本発明の目的は、複数台のディスクアレイ制御装置を１つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置間でのデータ移行による性能低下を抑えることが可能なディスクアレイシステムを提供すること、また、ディスクアレイ制御装置が有する機能を、性能低下を抑えて複数台のディスクアレイ制御装置で実現することにある。

【００２９】

【課題を解決するための手段】上記目的は、ホストコンピュータとのインターフェースを有する１または複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する１または複数のディスクインターフェース部と、前記ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記ディスク装置の管理情報を格納する共有メモリ部とを有し、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部が接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部が接続され、前記ホストコンピュータからのデータのリード／ライト要求に対し、前記チャネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うそれ自体ディスク制御装置としての機能を備えているディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置であって、前記複数のディスクアレイ制御ユニット内の前記共有メモリ部間を接続する手段と、前記複数のディスクアレイ制御ユニット内の前記キャッシュメモリ部間を接続する手段を有し、前記ディスクアレイ制御ユニット内の前記チャネルインターフェース部及び前記ディスクインターフェース部から、他の前記ディスクアレイ制御ユニット内の前記共有メモリ部のデータ、または前記キャッシュメモリ部のデータをリード／ライト可能であることを特徴とするディスクアレイ制御装置により達成される。

【００３０】

【発明の実施の形態】以下、大容量のデータの記憶装置として磁気ディスク装置を例にとって説明するが、大容量記憶装置として磁気ディスクに限られるものではなく、例えばＤＶＤのような大容量記憶装置であって良い。

【００３１】本発明の実施の形態の１つとして、好ましくは、前記複数のディスクアレイ制御ユニット内の前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数のキャッシュメモリ部との間は、前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続し、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続する。

【００３２】また、好ましくは、前記複数のディスクアレイ制御ユニット内の前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数のキャッシュメモリ部との間は、前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続し、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記ディスクアレイ制御ユニット内では直接接続し、前記ディスクアレイ制御ユニット間では、前記共有メモリ部間を前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続する。

【００３３】また、好ましくは、前記複数のディスクアレイ制御ユニット内の前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数のキャッシュメモリ部との間は、前記ディスクアレイ制御ユニット内では、自ユニット内のスイッチを用いた相互接続網によって接続し、前記複数のディスクアレイ制御ユニット間では、前記キャッシュメモリ部間を前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続し、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記複数の共有メモリ部との間は、前記ディスクアレイ制御ユニット内では、直接接続し、前記複数のディスクアレイ制御ユニット間では、前記共有メモリ部間を前記複数のディスクアレイ制御ユニット間に跨るスイッチを用いた相互結合網によって接続する。

【００３４】また、ホストコンピュータと磁気ディスク装置とのデータのリード／ライトに着目すれば以下のようである。ホストコンピュータとのインターフェースを有するチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有するディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報及び前記磁気ディスク装置の管理情報を格納する共有メモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部を接続する手段と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記共有メモリ部を接続する手段とを有し、前記ホストコンピュータからのデータのリード／ライト要

求に対し、前記チャネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、前記ディスクインターフェース部は、前記磁気ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行することにより、データのリード／ライトを行うディスクアレイ制御ユニットを、複数ユニット有するディスクアレイ制御装置であって、前記複数のディスクアレイ制御ユニット内の前記共有メモリ部間を接続する手段と、前記複数のディスクアレイ制御ユニット内の前記キャッシュメモリ部間を接続する手段を有し、該接続手段を介して、１つの前記ディスクアレイ制御ユニットにのみ接続されている前記ホストコンピュータから、該ディスクアレイ制御ユニットとは異なる前記ディスクアレイ制御ユニットにのみ接続されている前記磁気ディスク装置のデータをリード／ライト可能である。

【００３５】その他、本願が開示する課題、及びその解決方法は、発明の実施形態の欄及び図面により明らかにされる。

【００３６】以下、本発明の実施例を図面を用いて説明する。

【００３７】［実施例１］図１、図７、及び図８に、本発明の一実施例を示す。以下の実施例において、相互結合網はスイッチを利用したものを例に説明してあるが、相互に接続され制御情報やデータが転送されればよいのであり、例えばバスで構成されても良い。

【００３８】図１に示すように、ディスクアレイ制御装置１は複数のディスクアレイ制御ユニット１－１から構成される。ディスクアレイ制御ユニット１－１は、ホストコンピュータ５０とのインターフェース部（チャネルＩＦ部）１１と、磁気ディスク装置５とのインターフェース部（ディスクＩＦ部）１２と、共有メモリ部１３と、キャッシュメモリ部１４を有し、チャネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１３の間は複数のディスクアレイ制御ユニット１－１に跨る相互結合網２１０を介して接続され、チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４の間は複数のディスクアレイ制御ユニット１－１に跨る相互結合網２２０を介して接続されている。すなわち、相互結合網２１０、あるいは相互結合網２２０を介して、全てのチャネルＩＦ部１１及びディスクＩＦ部１２から、全ての共有メモリ部１３、あるいは全てのキャッシュメモリ部１４へアクセス可能な構成となっている。

【００３９】１つのディスクアレイ制御ユニットは１つの筐体として構成されるか、またはモジュールとして構成されても良いが、それ自体１つのディスク制御装置としての機能を備えているものである。図７ではそれぞれ別個な筐体を構成しているものとして説明する。

【００４０】１つのディスクアレイ制御ユニット１－１内の具体的な一例を図７に示す。ディスクアレイ制御ユニット１－１は、ホストコンピュータ５０との２つのチャネルＩＦ部１１と、磁気ディスク装置５との２つのディスクＩＦ部１２と、２つの共有メモリパス用スイッチ（ＳＭ－ＳＷ）１１０と、２つのキャッシュメモリ用パススイッチ（ＣＭ－ＳＷ）１１１と、２つの共有メモリ部１３と、２つのキャッシュメモリ部１４と、共有メモリ（ＳＭ）アクセスパス１３５と、１３６と、キャッシュメモリ（ＣＭ）アクセスパス１３７と、１３８と、筐体間ＳＭパス１４１と、筐体間ＣＭパス１４２を有する。筐体間ＳＭパスはディスクアレイ制御ユニット間ＳＭパスと、筐体間ＣＭパスはディスクアレイユニット間ＣＭパスと言うことが出きる。

【００４１】チャネルＩＦ部１１は、ホストコンピュータ５０との２つのＩＦ（ホストＩＦ）１０２と、ホストコンピュータ５０に対する入出力を制御する２つのマイクロプロセッサ１０１と、共有メモリ部１３へのアクセスを制御するアクセス制御部（ＳＭアクセス制御部）１０４と、キャッシュメモリ部１４へのアクセスを制御するアクセス制御部（ＣＭアクセス制御部）１０５を有し、ホストコンピュータ５０とキャッシュメモリ部１４間のデータ転送、及びマイクロプロセッサ１０１と共有メモリ部１３間の制御情報の転送を実行する。マイクロプロセッサ１０１及びホストＩＦ１０２は内部バス１０６によって接続され、ＣＭアクセス制御部１０５は２つのホストＩＦ１０２に直接接続されている。また、ＳＭアクセス制御部１０４は２つのマイクロプロセッサ１０１に直接接続されている。

【００４２】ディスクＩＦ部１２は、磁気ディスク装置５との２つのＩＦ（ドライブＩＦ）１０３と、磁気ディスク装置５に対する入出力を制御する２つのマイクロプロセッサ１０１と、共有メモリ部１３への１つのアクセス制御部（ＳＭアクセス制御部）１０４と、キャッシュメモリ部１４への１つのアクセス制御部（ＣＭアクセス制御部）１０５を有し、磁気ディスク装置５とキャッシュメモリ部１４間のデータ転送、及びマイクロプロセッサ１０１と共有メモリ部１３間の制御情報の転送を実行する。マイクロプロセッサ１０１及びドライブＩＦ１０３は内部バス１０６によって接続され、ＣＭアクセス制御部１０５は２つのドライブＩＦ１０３に直接接続されている。また、ＳＭアクセス制御部１０４は２つのマイクロプロセッサ１０１に直接接続されている。ディスクＩＦ部はRAID機能の実行も行う。

【００４３】共有メモリ部１３は、共有メモリ（ＳＭ）コントローラ１０７とメモリモジュール１０９とを有し、ディスクアレイ制御ユニット１－１の制御情報（例えば、チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４との間のデータ転送制御に関する情報、磁気ディスク装置５に記録するデータの管理情報）等を格納する。

【００４４】キャッシュメモリ部１４は、キャッシュメ

モリ（ＣＭ）コントローラ１０８とメモリモジュール１０９を有し、磁気ディスク装置５へ記録するデータを一時的に格納する。

【００４５】ＳＭアクセス制御部１０４には２本のＳＭアクセスパス１３５を接続し、それらを２つの異なるＳＭ－ＳＷ１１０にそれぞれ接続する。ＳＭ－ＳＷ１１０には２本のアクセスパス１３６を接続し、それらを２つの異なるＳＭコントローラ１０７にそれぞれ接続する。したがってＳＭコントローラ１０７には、２つのＳＭ－ＳＷ１１０から１本ずつ、計２本のアクセスパス１３６が接続される。こうすることにより、１つのＳＭアクセス制御部１０４から１つのＳＭコントローラ１０７へのアクセスルートが２つとなる。これにより、１つのアクセスパスまたはＳＭ－ＳＷ１１０に障害が発生した場合でも、もう１つのアクセスルートにより共有メモリ部１３へアクセスすることが可能となるため、耐障害性を向上させることができる。

【００４６】ＣＭアクセス制御部１０５には２本のＣＭアクセスパス１３７を接続し、それらを２つの異なるＣＭ－ＳＷ１１１にそれぞれ接続する。ＣＭ－ＳＷ１１１には２本のアクセスパス１３８を接続し、それらを２つの異なるＣＭコントローラ１０８にそれぞれ接続する。したがってＣＭコントローラ１０８には、２つのＣＭ－ＳＷ１１１から１本ずつ、計２本のアクセスパス１３８が接続される。こうすることにより、１つのＣＭアクセス制御部１０５から１つのＣＭコントローラ１０８へのアクセスルートが２つとなる。これにより、１つのアクセスパスまたはＣＭ－ＳＷ１１１に障害が発生した場合でも、もう１つのアクセスルートによりキャッシュメモリ部１４へアクセスすることが可能となるため、耐障害性を向上させることができる。

【００４７】ＳＭ－ＳＷ１１０には、２つのチャネルＩＦ部１１と、２つのディスクＩＦ部１２からそれぞれ１本ずつ、計４本のＳＭアクセスパス１３５が接続される。また、ＳＭ－ＳＷ１１０には、２つの共有メモリ部１３へのアクセスパス１３６が１本ずつ、計２本接続される。また、他のディスクアレイ制御ユニット１－１のＳＭ－ＳＷ１１０と接続するための筐体間ＳＭパス１４１が２本接続される。これは一方が入力、他方が出力用であっても良く、また、それぞれが双方向の情報の転送が出来るものであっても良い。

【００４８】ＣＭ－ＳＷ１１１には、２つのチャネルＩＦ部１１と、２つのディスクＩＦ部１２からそれぞれ１本ずつ、計４本のＣＭアクセスパス１３７が接続される。また、ＣＭ－ＳＷ１１１には、２つのキャッシュメモリ部１４へのアクセスパス１３８が１本ずつ、計２本接続される。また、他のディスクアレイ制御ユニット１－１のＣＭ－ＳＷ１１１と接続するための筐体間ＣＭパス１４２が２本接続される。この２本もまた筐体間ＳＭパスと同様の性質を持っている。

【００４９】ＳＭ－ＳＷ１１０、あるいはＣＭ－ＳＷ１１１では、上記のようなアクセスパスが接続されるため、ＳＭ－ＳＷ１１０内、あるいはＣＭ－ＳＷ１１１内では、チャネルＩＦ部１１及びディスクＩＦ部１２からの４本のアクセスパスからの要求を、自ディスクアレイ制御ユニット１－１内の共有メモリ部１３、あるいはキャッシュメモリ部１４への２本のアクセスパスと、他ディスクアレイ制御ユニット１－１内の共有メモリ部１３、あるいはキャッシュメモリ部１４への２本の筐体間アクセスパスに振分ける機能を有する。

【００５０】図７でＳＭ－ＳＷ１１０はチャネルＩＦ部１１およびディスクＩＦ部１２と共有メモリ部１３との接続部であり、ＣＭ－ＳＷ１１１はチャネルＩＦ部１１およびディスクＩＦ部１２とキャッシュメモリ部１４との接続部である。

【００５１】図７に示すディスクアレイ制御ユニット１－１を２台接続したディスクアレイ制御装置１の例を図８に示す。

【００５２】各ディスクアレイ制御ユニット１－１－１、１－１－２内のＳＭ－ＳＷ１１０間は、各筐体間ＳＭパス１４１を筐体間ＳＭ－ＳＷ１２１を介して接続する。

【００５３】同様に、各ディスクアレイ制御ユニット１－１－１、１－１－２内のＣＭ－ＳＷ１１１間は、各筐体間ＣＭパス１４２を筐体間ＣＭ－ＳＷ１２２を介して接続する。ＳＷ１２１、１２２はスイッチボックスとして実装される。

【００５４】本実施例のように、接続するディスクアレイ制御ユニット１－１が２台の場合は、筐体間ＳＷ１２１、１２２を介さずに直接筐体間パスを接続しても本発明を実施する上で問題は無い。そうすることにより、筐体間ＳＷ１２１、１２２で発生するデータ転送処理のオーバヘッドを削減することが可能となり、性能が向上する。

【００５５】筐体間ＳＷ１２１、１２２に接続できるパス数を、図８に示す例より増やすことにより、３台以上のディスクアレイ制御ユニット１－１を接続することが可能となる。また、筐体間ＳＷ１２１、１２２に実装可能なパス数には物理的な限界があるため、接続するディスクアレイ制御ユニット１－１を増やしていった場合、１つの筐体間ＳＷだけでは接続できなくなる。そうした場合は、筐体間ＳＷを多段に接続することで、接続可能なディスクアレイ制御ユニット１－１の台数を増やすことが可能となる。スイッチボックスにより３台のディスクアレイ制御ユニットを相互結合した実装例を図１９に示す。詳細は後述する。

【００５６】また、３台以上のディスクアレイ制御ユニット１－１を接続する場合、図１３に示すように各ディスクアレイ制御ユニット１－１内のＳＭ－ＳＷ１１０及びＣＭ－ＳＷ１１１間をループ状に接続する。このよう

な構成では、筐体間ＳＷ１２１、１２２を用いずに複数のディスクアレイ制御ユニット１－１を接続することが可能となる。

【００５７】この場合には各ディスクアレイ制御ユニット１－１から出ている筐体間ＳＭパス同士、筐体間ＣＭパス同士をコネクタで接続すれば良い。図では省略してあるが筐体間ＳＭパス１４１、および筐体間ＣＭパス１４２のディスクアレイ制御ユニット１－１の外の部分にコネクタを設ければディスクアレイ制御ユニットを増設するのに都合が良い。図８でＳＭ－ＳＷ１１０、筐体間ＳＭパス１４１および筐体間ＳＭ－ＳＷ１２１で図１の相互結合網２１０を構成し、ＣＭ－ＳＷ１１１、筐体間ＣＭパス１４２および筐体間ＣＭ－ＳＷ１２２で図１の相互結合網２２０を構成する。

【００５８】図８において、ディスクアレイ制御装置１－１－１に接続されたホストコンピュータ５０からデータを読み出す場合の一例を述べる。

【００５９】まず、ホストコンピュータ５０は、自身が接続されているディスクアレイ制御装置１－１－１内のチャネルＩＦ部１１にデータの読出し要求を発行する。要求を受けたチャネルＩＦ部１１内のマイクロプロセッサ１０１は、２つのディスクアレイ制御ユニット１－１－１、１－１－２内の共有メモリ部１３の両方にアクセスし、要求されたデータがどの磁気ディスク装置５内に格納されているかを調べる。共有メモリ部１３には、要求データのアドレスとそのデータがどの磁気ディスク装置５に格納されているかを対応させる変換テーブルが格納されており、要求されたデータがどの磁気ディスク装置５内に格納されているかを調べることができる。

【００６０】次に、要求を受けたチャネルＩＦ部１１内のマイクロプロセッサ１０１は、２つのディスクアレイ制御ユニット１－１－１、１－１－２内の共有メモリ部１３の両方にアクセスし、各ディスクアレイ制御ユニット１－１－１、１－１－２内のキャッシュメモリ部１４に要求されたデータが格納されているかどうかを確認する。共有メモリ部１３にはキャッシュメモリ部１４内データのディレクトリ情報が格納されており、各キャッシュメモリ部１４に要求データが存在するかどうかを確認できる。

【００６１】それによりディスクアレイ制御ユニット１－１－１のキャッシュメモリ部１４内にデータがあった場合は、そのデータを自身のＣＭ－ＳＷ１１１を介してチャネルＩＦ部１１まで転送し、ホストコンピュータ５０に送る。ディスクアレイ制御ユニット１－１－２のキャッシュメモリ部１４内にデータがあった場合は、そのデータをディスクアレイ制御ユニット１－１－２内のＣＭ－ＳＷ１１１、筐体間ＣＭ－ＳＷ１２２、自身のＣＭ－ＳＷ１１１を介してチャネルＩＦ部１１まで転送し、ホストコンピュータ５０に送る。

【００６２】どのキャッシュメモリ部１４内にもデータが存在しなかった場合、マイクロプロセッサ１０１は要求データが格納されている磁気ディスク装置５が接続されているディスクＩＦ部１２内のマイクロプロセッサ１０１に対し、要求データを読出し、キャッシュメモリ部１４に格納するように命令を発行する。命令を受けたディスクＩＦ部１２内のマイクロプロセッサ１０１は、要求データが格納されている磁気ディスク装置５からデータを読出し、２つのディスクアレイ制御ユニット１－１－１、１－１－２の内のどちらか一方のキャッシュメモリ部１４に要求データを格納する。要求データが格納されている磁気ディスク装置５が繋がるディスクアレイ制御ユニット１－１－２内のキャッシュメモリ部１４にデータを格納する場合は、そのディスクアレイ制御ユニット１－１－２内のＣＭ－ＳＷ１１１を介することにより、また、要求データが格納されている磁気ディスク装置５が繋がるディスクアレイ制御ユニット１－１－２とは異なるディスクアレイ制御ユニット１－１－１内のキャッシュメモリ部１４にデータを格納する場合は、それぞれのＣＭ－ＳＷ１１１及び筐体間ＣＭ－ＳＷ１２２を介することにより、キャッシュメモリ部１４へデータが転送される。

【００６３】ディスクＩＦ部１２内のマイクロプロセッサ１０１は、要求データをキャッシュメモリ部１４へ格納した後、命令を発行したチャネルＩＦ部１１内のマイクロプロセッサ１０１に、データを格納したキャッシュメモリ部１４を伝える。それを受けたチャネルＩＦ部１１内のマイクロプロセッサ１０１は、データが格納されたキャッシュメモリ部１４からデータを読出し、ホストコンピュータ５０へ送る。

【００６４】本実施例によれば、ホストコンピュータ５０は要求データがどのディスクアレイ制御ユニット１－１に繋がる磁気ディスク装置５に格納されているかを意識することなく、自身が繋がるディスクアレイ制御ユニット１－１にアクセス要求を発行するだけで、データの書き込み及び読出しを行うことが可能になり、ホストコンピュータ５０に対して、複数台のディスクアレイ制御ユニット１－１を１つのディスクアレイ制御装置１に見せることが可能となる。

【００６５】また、要求を受けたディスク制御ユニット１－１とは異なるディスクアレイ制御ユニット１－１に繋がる磁気ディスク装置５からデータを読み出す場合、内部の相互結合網及びキャッシュメモリ部１４を介して、データを読み出すことが出来るため、両方のディスクアレイ制御ユニット１－１のチャネルＩＦ部１１を介してデータを移行する必要が無くなり、データの読出し及び書き込みの性能の低下を抑えることが可能となる。

【００６６】［実施例２］図５、図９、及び図１０に、本発明の他の実施例を示す。

【００６７】図５に示すように、複数のディスクアレイ制御ユニット１－２からなるディスクアレイ制御装置１

の構成は、チャネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１３の間の接続構成を除いて、実施例１の図１に示す構成と同様である。

【００６８】チャネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１３の間は、ディスクアレイ制御ユニット１―２内において直接に接続されている。また、複数のディスクアレイ制御ユニット１－２の間では、共有メモリ部１３は相互結合網２４を介して接続されている。

【００６９】上記のように、この実施例ではディスクアレイ制御ユニット１－２内においてチャネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１３を直接接続することにより、実施例１で示した相互結合網２１０を介して接続する場合に比べ、共有メモリ部１３へのアクセス時間を短縮することが可能になる。

【００７０】１つのディスクアレイ制御ユニット１－２内の具体的な一例を図９に示す。

【００７１】ディスクアレイ制御ユニット１－２内の構成も、チャネルＩＦ部１１及びディスクＩＦ部１２と共有メモリ部１３の間の接続構成を除いて、実施例１の図７に示す構成と同様である。

【００７２】ディスクアレイ制御ユニット１－２は、ホストコンピュータ５０との２つのチャネルＩＦ部１１と、磁気ディスク装置５との２つのディスクＩＦ部１２と、２つのキャッシュメモリ用パススイッチ（ＣＭ－ＳＷ）１１１と、２つの共有メモリ部１３と、２つのキャッシュメモリ部１４と、共有メモリ（ＳＭ）アクセスパス１３９と、キャッシュメモリ（ＣＭ）アクセスパス１３７と、ＣＭアクセスパス１３８と、筐体間ＳＭパス１４３と、筐体間ＣＭパス１４２を有する。

【００７３】ＳＭアクセス制御部１０４には２本のＳＭアクセスパス１３９を接続し、それらを２つの異なるＳＭコントローラ１０７にそれぞれ接続する。したがってＳＭコントローラ１０７には、２つのチャネルＩＦ部１１及び２つのディスクＩＦ部１２から１本ずつ、計４本のＳＭアクセスパス１３９が接続される。また、他のディスクアレイ制御ユニット１－２のＳＭコントローラ１０７と接続するための筐体間ＳＭパス１４３が２本接続される。

【００７４】ＳＭコントローラ１０７では、上記のようなアクセスパスが接続されるため、ＳＭコントローラ１０７内では、チャネルＩＦ部１１及びディスクＩＦ部１２からの４本のＳＭアクセスパス１３９からの要求を、メモリモジュール１０９へのアクセスパスと、他ディスクアレイ制御ユニット１―２内の共有メモリ部１３への２本の筐体間ＳＭアクセスパス１４３に振分ける機能を有する。

【００７５】図９に示すディスクアレイ制御ユニット１－２を２台接続したディスクアレイ制御装置１の例を図１０に示す。

【００７６】各ディスクアレイ制御ユニット１－２内の共有メモリ部１３間は、各筐体間ＳＭパス１４３を筐体間ＳＭ－ＳＷ１２１を介して接続する。

【００７７】それ以外は、実施例１の図８に示す構成と同様である。この場合はＳＭコントローラ１０７がチャネルＩＦ部とディスクＩＦ部と共有メモリ部との接続部をなす。

【００７８】実施例１と同様に、接続するディスクアレイ制御ユニット１－２が２台の場合は、筐体間ＳＷ１２１、１２２を介さずに直接筐体間パスを接続しても本発明を実施する上で問題は無い。そうすることにより、筐体間ＳＷ１２１、１２２で発生するデータ転送処理のオーバヘッドを削減することが可能となり、性能が向上する。

【００７９】また、実施例１と同様に、筐体間ＳＷ１２１、１２２に接続できるパス数を、図１０に示す例より増やすことにより、３台以上のディスクアレイ制御ユニット１－２を接続することが可能となる。また、筐体間ＳＷ１２１、１２２に実装可能なパス数には物理的な限界があるため、接続するディスクアレイ制御ユニット１－２を増やしていった場合、１つの筐体間ＳＷだけでは接続できなくなる。そうした場合は、筐体間ＳＷを多段に接続することで、接続可能なディスクアレイ制御ユニット１－２の台数を増やすことが可能となる。これは実施例１と同様スイッチボックスとして実装される。

【００８０】また、３台以上のディスクアレイ制御ユニット１－２を接続する場合、実施例１において、図１３に示した構成と同様のループ状の構成をとることにより、、筐体間ＳＷ１２１、１２２を用いずに複数のディスクアレイ制御ユニット１－２を接続することが可能となる。

【００８１】本実施例において、ホストコンピュータ５０から磁気ディスク装置５へのデータの読み出し／書き込みを行う場合の、ディスクアレイ制御ユニット１－２内での各部の動作は、他のディスクアレイ制御ユニット１－２内の共有メモリ部１３へアクセスが、自ディスクアレイ制御ユニット１－２内の共有メモリ部１３及び相互結合網２４を介して行われることを除いて、実施例１と同様である。なお、筐体間ＳＭパス１４３、筐体間ＳＭ－ＳＷ１２１、およびＳＭコントローラ１０７で相互結合網２４が構成される。

【００８２】本実施例によれば、ホストコンピュータ５０は要求データがどのディスクアレイ制御ユニット１－２に繋がる磁気ディスク装置５に格納されているかを意識することなく、自身が繋がるディスクアレイ制御ユニット１－２にアクセス要求を発行するだけで、データの書き込み及び読出しを行うことが可能になり、ホストコンピュータ５０に対して、複数台のディスクアレイ制御ユニット１－２を１つのディスクアレイ制御装置１に見せることが可能となる。

【００８３】また、要求を受けたディスク制御ユニット１－２とは異なるディスクアレイ制御ユニット１－２に繋がる磁気ディスク装置５からデータを読み出す場合、内部の相互結合網及びキャッシュメモリ部１４を介して、データを読み出すことが出来るため、両方のディスクアレイ制御ユニット１－２のチャネルＩＦ部１１を介してデータを移行する必要が無くなり、データの読出し及び書き込みの性能の低下を抑えることが可能となる。

【００８４】［実施例３］図６、図１１、及び図１２に、本発明の他の実施例を示す。

【００８５】図６に示すように、複数のディスクアレイ制御装置１－３からなるディスクアレイ制御装置１の構成は、チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４の間の接続構成を除いて、実施例２の図５に示す構成と同様である。

【００８６】チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４の間は、ディスクアレイ制御ユニット１－３内において相互結合網２２を介して接続されている。また、複数のディスクアレイ制御ユニット１－３の間では、キャッシュメモリ部１４は相互結合網２５を介して接続されている。実施例２と同様にチャネルＩＦ部１１およびディスクＩＦ部１２と共有メモリ部１３が直接接続されているのに対し、チャネルＩＦ部１１およびディスクＩＦ部１２とキャッシュメモリ部１４との間は相互結合網２２を介して接続されているのは次の理由である。共有メモリ部１３で扱われる制御情報は例えば数バイトであるのに対し、キャッシュメモリ部１４では例えば数Ｋバイトの単位でデータが扱われる。そのため、相互結合網２２を通して限られたピン数で接続しながらスループットを上げることが図られる。

【００８７】上記のように、チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４を接続する相互結合網２２と、ディスクアレイ制御ユニット１－３間でキャッシュメモリ部１４を接続する相互結合網２４を分離することにより、異なるディスクアレイ制御ユニット１－３のキャッシュメモリ部１４間でデータを移行させる場合、ホストコンピュータ５０からのアクセス要求時のキャッシュメモリ部１４へのアクセスを妨げることがなくなる。キャッシュメモリ間のデータの移行はディスクＩＦ部が司る。異なるディスクアレイ制御ユニット１－３のキャッシュメモリ部１４間でデータを移行させる機能は、１つのディスクアレイ制御ユニット１－３に対してホストコンピュータ５０からのアクセスが集中した場合、アクセス頻度の少ないディスクアレイ制御ユニットにデータを移動させるために必要な機能である。

【００８８】１つのディスクアレイ制御ユニット１－３内の具体的な一例を図１１に示す。

【００８９】ディスクアレイ制御ユニット１－３内の構成も、チャネルＩＦ部１１及びディスクＩＦ部１２とキャッシュメモリ部１４の間の接続構成を除いて、実施例２の図９に示す構成と同様である。ここではＣＭコントローラ１０８がチャネルＩＦ部とディスクＩＦ部とキャッシュメモリ部との接続部をなす。

【００９０】ディスクアレイ制御ユニット１－３は、ホストコンピュータ５０との２つのチャネルＩＦ部１１と、磁気ディスク装置５との２つのディスクＩＦ部１２と、２つのキャッシュメモリ用パススイッチ（ＣＭ－ＳＷ）１１１と、２つの共有メモリ部１３と、２つのキャッシュメモリ部１４と、共有メモリ（ＳＭ）アクセスパス１３９と、キャッシュメモリ（ＣＭ）アクセスパス１３７と、ＣＭアクセスパス１３８と、筐体間ＳＭパス１４３と、筐体間ＣＭパス１４４を有する。

【００９１】ＣＭアクセス制御部１０５には２本のＣＭアクセスパス１３７を接続し、それらを２つの異なるＣＭ－ＳＷ１１１にそれぞれ接続する。ＣＭ－ＳＷ１１１には２本のアクセスパス１３８を接続し、それらを２つの異なるＣＭコントローラ１０８にそれぞれ接続する。したがってＣＭコントローラ１０８には、２つのＣＭ－ＳＷ１１１から１本ずつ、計２本のアクセスパス１３８が接続される。また、他のディスクアレイ制御ユニット１－３のＣＭコントローラ１０８と接続するための筐体間ＣＭパス１４４が２本接続される。

【００９２】ＣＭコントローラ１０８では、上記のようなアクセスパスが接続されるため、ＣＭコントローラ１０８内では、ＣＭ－ＳＷ１１１からの２本のＣＭアクセスパス１３８からの要求を、メモリモジュール１０９へのアクセスパスと、他ディスクアレイ制御ユニット１－３内のキャッシュメモリ部１４への２本の筐体間ＣＭアクセスパス１４４に振分ける機能を有する。

【００９３】ＣＭ－ＳＷ１１１には、２つのチャネルＩＦ部１１と、２つのディスクＩＦ部１２からそれぞれ１本ずつ、計４本のＣＭアクセスパス１３７が接続される。また、ＣＭ－ＳＷ１１１には、２つのキャッシュメモリ部１４へのアクセスパス１３８が１本ずつ、計２本接続される。

【００９４】ＣＭ－ＳＷ１１１では、上記のようなアクセスパスが接続されるため、ＣＭ－ＳＷ１１１内では、チャネルＩＦ部１１及びディスクＩＦ部１２からの４本のＣＭアクセスパス１３７からの要求を、キャッシュメモリ部１４への２本のＣＭアクセスパス１３８に振分ける機能を有する。

【００９５】図１１に示すディスクアレイ制御ユニット１－３を２台接続したディスクアレイ制御装置１の例を図１２に示す。

【００９６】各ディスクアレイ制御ユニット１－３内のキャッシュメモリ部１４間は、各筐体間ＣＭパス１４４を筐体間ＣＭ－ＳＷ１２２を介して接続する。それ以外は、実施例２の図１０に示す構成と同様である。

【００９７】実施例２と同様に、接続するディスクアレイ制御ユニット１－３が２台の場合は、筐体間ＳＷ１２

１、１２２を介さずに直接筐体間パスを接続しても本発明を実施する上で問題は無い。そうすることにより、筐体間ＳＷ１２１、１２２で発生するデータ転送処理のオーバヘッドを削減することが可能となり、性能が向上する。

【００９８】また、実施例２と同様に、筐体間ＳＷに接続できるパス数を、図１２に示す例より増やすことにより、３台以上のディスクアレイ制御ユニット１－３を接続することが可能となる。また、筐体間ＳＷ１２１、１２２に実装可能なパス数には物理的な限界があるため、接続するディスクアレイ制御ユニット１－３を増やしていった場合、１つの筐体間ＳＷだけでは接続できなくなる。そうした場合は、筐体間ＳＷを多段に接続することで、接続可能なディスクアレイ制御ユニット１２１、１２２の台数を増やすことが可能となる。

【００９９】また、３台以上のディスクアレイ制御ユニット１－３を接続する場合、実施例１において、図１３に示した構成と同様のループ状の構成をとることにより、筐体間ＳＷ１２１、１２２を用いずに複数のディスクアレイ制御ユニット１－３を接続することが可能となる。

【０１００】本実施例において、ホストコンピュータ５０から磁気ディスク装置５へのデータの読み出し／書き込みを行う場合の、ディスクアレイ制御ユニット１－３内での各部の動作は、他のディスクアレイ制御ユニット１－３内のキャッシュメモリ部１４へアクセスが、自ディスクアレイ制御ユニット１－３内のキャッシュメモリ部１４及び相互結合網２５を介して行われることを除いて、実施例２と同様である。

【０１０１】本実施例によれば、ホストコンピュータ５０は要求データがどのディスクアレイ制御ユニット１－３に繋がる磁気ディスク装置５に格納されているかを意識することなく、自身が繋がるディスクアレイ制御ユニット１－３にアクセス要求を発行するだけで、データの書き込み及び読出しを行うことが可能になり、ホストコンピュータ５０に対して、複数台のディスクアレイ制御ユニット１－３を１つのディスクアレイ制御装置１に見せることが可能となる。

【０１０２】また、要求を受けたディスク制御ユニット１－３とは異なるディスクアレイ制御ユニット１－３に繋がる磁気ディスク装置５からデータを読み出す場合、内部の相互結合網及びキャッシュメモリ部１４を介して、データを読み出すことが出来るため、両方のディスクアレイ制御ユニット１－３のチャネルＩＦ部１１を介してデータを移行する必要が無くなり、データの読出し及び書き込みの性能の低下を抑えることが可能となる。

【０１０３】次に本発明の実施例の利用例について説明する。ハイエンドのディスクアレイ制御装置には以下のような機能がある。すなわち、ある業務用のデータセット(論理ボリュームに対応)の複製を保持しておき、通常の業務ではオリジナルデータセットと複製データセットに対して同時にデータの更新を行い、例えば、そのデータセットのバックアップをとる要求があった場合、複製データセットについてデータの更新を中止し、それをバックアップアップ用に使用し、オリジナルデータセットは業務を継続し、バックアップが終了した時点で、オリジナルデータセットと複製のデータセットのデータの整合性をとるという機能がある。

【０１０４】実施例１に示すディスクアレイ制御装置１において、異なるディスクアレイ制御ユニット１－１間でデータセットの複製を保持する場合に、上記機能を実現する方法について、図８を用いて説明する。

【０１０５】ここでは仮に、図８に示すディスクアレイ制御ユニット１－１－１に繋がる磁気ディスク装置５にオリジナルのデータセットが格納され、ディスクアレイ制御ユニット１－１－２に繋がる磁気ディスク装置５に複製のデータセットが格納されているとする。また、ディスクアレイ制御ユニット１－１－１に繋がるホストコンピュータ５０が通常の業務を行い、ディスクアレイ制御ユニット１－１－２に繋がるホストコンピュータ５０が自身に繋がるテープ装置(図示していない)にデータのバックアップをとる作業を行うことにする。

【０１０６】通常の業務では、ディスクアレイ制御ユニット１－１－１に繋がるホストコンピュータ５０から目的のデータセットへデータの書き込み要求があった場合、ディスクアレイ制御ユニット１－１－１に繋がるホストコンピュータ５０に繋がるチャネルＩＦ部１１内のマイクロプロセッサ１０１は、ホストコンピュータ５０から送られてきたデータをディスクアレイ制御ユニット１－１－１のキャッシュメモリ部１４に転送し、書き込む。次に、前記のマイクロプロセッサ１０１は、オリジナルのデータセットが格納されている磁気ディスク装置５が繋がるディスクＩＦ部１２内のマイクロプロセッサ１０１に、ディスクアレイ制御ユニット１－１－１の共有メモリ部１３を介して命令を発行し、ディスクアレイ制御ユニット１－１－１のキャッシュメモリ部１４からデータを読み出し、オリジナルのデータセットが格納されている磁気ディスク装置５に繋がるディスクＩＦ部１２へ転送し、そこから磁気ディスク装置５に転送し、書き込ませる。

【０１０７】ディスクアレイ制御ユニット１－１－１のチャネルＩＦ部１１内のマイクロプロセッサ１０１は、オリジナルのデータセットのデータ更新を監視しており、ディスクアレイ制御ユニット１－１－１の共有メモリ部１３内にオリジナルのデータセットのデータ更新量を示す制御情報を格納する。データ更新量があらかじめ定められた値以上になると、前記のマイクロプロセッサ１０１は、オリジナルのデータセットの更新内容を複製のデータセットに反映するように、オリジナルのデータセットが格納されている磁気ディスク装置５が繋がるデ

ィスクＩＦ部１２内のマイクロプロセッサ１０１に命令を発行する。それを受けて、マイクロプロセッサ１０１は、磁気ディスク装置５から更新されたデータを読み出し、その更新データのアドレスを複製のデータセットのアドレスに変換する。その更新データを、ディスクアレイ制御ユニット１－１－１のＣＭ－ＳＷ１１１、筐体間ＣＭ－ＳＷ１２２、及びディスクアレイ制御ユニット１－１－２のＣＭ－ＳＷ１１１を介して、ディスクアレイ制御ユニット１－１－２のキャッシュメモリ部１４へ転送し、書き込む。次に、そのキャッシュメモリ部１４から更新データを読み出し、複製のデータセットが格納されている磁気ディスク装置５に繋がるディスクＩＦ部１２へ転送し、そこから磁気ディスク装置５に転送し、書きこむ。

【０１０８】上記の動作により、通常業務において、オリジナルのデータセットと複製のデータセットを保持する。

【０１０９】ディスクアレイ制御ユニット１－１－２に繋がるホストコンピュータ５０から、目的のデータセットに対してバックアップをとる要求があった場合、そのホストコンピュータ５０が繋がるチャネルＩＦ部１１内のマイクロプロセッサ１０１は、ディスクアレイ制御ユニット１－１－１の共有メモリ部１３を介して、通常業務を行っているホストコンピュータ５０が繋がるチャネルＩＦ部１１内のマイクロプロセッサ１０１に複製のデータセットに対するデータの更新を一時中断するように命令を発行する。それを受けたマイクロプロセッサ１０１は、データの更新を一時中断する。次に、バックアップの要求を出したホストコンピュータ５０に繋がるチャネルＩＦ部１１内のマイクロプロセッサ１０１は、複製のデータセットが格納されている磁気ディスク装置５に繋がるディスクＩＦ部１２内のマイクロプロセッサ１０１に、ディスクアレイ制御ユニット１－１－２の共有メモリ部１３を介して命令を発行し、複製のデータセットを磁気ディスク装置５から読み出し、ディスクＩＦ部１２へ転送し、ディスクＩＦ部１２からディスクアレイ制御ユニット１－１－２のキャッシュメモリ部１４に転送し、書き込ませる。それが終了すると、前記のチャネルＩＦ部１１内のマイクロプロセッサ１０１は、ディスクアレイ制御ユニット１－１－２のキャッシュメモリ部１４から複製のデータセットを読み出し、チャネルＩＦ部１１へ転送し、そこからバックアップの要求を出したホストコンピュータ５０へ送る。

【０１１０】データセットのバックアップが終了すると、バックアップの要求を出したホストコンピュータ５０に繋がるチャネルＩＦ部１１内のマイクロプロセッサ１０１は、通常業務を行うホストコンピュータ５０が繋がるチャネルＩＦ部１１内のマイクロプロセッサ１０１に、ディスクアレイ制御ユニット１－１－１の共有メモリ部１３を介して命令を発行し、オリジナルのデータセット内のバックアップ処理中に更新されたデータを、複製のデータセットに反映させる。この方法は、上記の通常業務における更新データの反映を行う方法と同様である。本例によれば、上記機能を実現する場合、内部の相互結合網及びキャッシュメモリ部１４を介して、２つのディスクアレイ制御ユニット１－１－１、１－１－２間でデータの移行を行えるため、両方のディスクアレイ制御ユニット１－１－１、１－１－２のチャネルＩＦ部を介してデータを移行する必要が無くなり、上記機能を実行しているときの性能低下を抑えることが可能となる。したがって、ユーザの通常業務の効率を低下させることがなくなる。

【０１１１】実施例２及び実施例３の構成のディスクアレイ制御装置１においても、本実施例を実施する上で問題はなく、本実施例と同様の効果が得られる。

【０１１２】その他の利用例として以下のものがある。実施例１、実施例２、及び実施例３に示すディスクアレイ制御装置１では、ホストコンピュータ５０からディスクアレイ制御ユニットに、例えばデータのリード要求があり、そのディスクアレイ制御ユニットに接続された磁気ディスク装置５にデータが無かった場合、相互結合網を介して、該当するデータを格納している磁気ディスク装置５が接続されたディスクアレイ制御ユニットからデータを読み出し、ホストコンピュータ５０へ要求データを返す必要がある。このようにディスクアレイ制御ユニット間を跨ってデータを移行させることによりデータのリード／ライトを行う場合は、そうでない場合に比べて性能が低下する。

【０１１３】上記のようなデータ移行を抑えるため、あるホストコンピュータ５０からアクセスされる頻度の高いデータの中で、該ホストコンピュータ５０が繋がるディスクアレイ制御ユニットとは異なるディスクアレイ制御ユニットに接続された磁気ディスク装置５に格納されているデータを、該ホストコンピュータ５０が繋がるディスクアレイ制御ユニットに接続された磁気ディスク装置５に移行する機能が必要となる。

【０１１４】実施例１に示すディスクアレイ制御装置１において、上記のデータ移行の方法を図８を用いて説明する。

【０１１５】チャネルＩＦ部１１内のマイクロプロセッサ１０１は、全ての磁気ディスク装置５内のデータセット(論理ボリュームに対応)へのアクセス頻度を監視しており、自身と同じディスクアレイ制御ユニット１－１－１内の共有メモリ部１３内に前記のデータセットのアクセス頻度を示す制御情報を格納する。

【０１１６】ここで、ディスクアレイ制御ユニット１－１－１に繋がるホストコンピュータ５０から、ディスクアレイ制御ユニット１－１－２に繋がる磁気ディスク装置５内のデータセットにアクセスが集中し、アクセス頻度があらかじめ定められた値を超えた場合、ディスクア

レイ制御ユニット１－１－１のチャネルＩＦ部１１内のマイクロプロセッサ１０１は、該当するデータセットを格納している磁気ディスク装置５が繋がるディスクＩＦ部１２内のマイクロプロセッサ１０１に、ディスクアレイ制御ユニット１－１－２内の共有メモリ部１３を介して命令を発行し、該当するデータセットを読み出し、ディスクアレイ制御ユニット１－１－２のキャッシュメモリ部１４へ転送し、書き込ませる。

【０１１７】次に、ディスクアレイ制御ユニット１－１－１のチャネルＩＦ部１１内のマイクロプロセッサ１０１は、ディスクアレイ制御ユニット１―１－２のキャッシュメモリ部１４から該当するデータを読み出し、ディスクアレイ制御ユニット１－１－１のキャッシュメモリ部１４に転送する。次に、ディスクアレイ制御ユニット１－１－１のディスクＩＦ部１２内のマイクロプロセッサ１０１に、ディスクアレイ制御ユニット１－１－１内の共有メモリ部１３を介して命令を発行し、ディスクアレイ制御ユニット１―１－１のキャッシュメモリ部１４から該当するデータを読み出し、磁気ディスク装置５に書き込ませる。本例によれば、上記のようなデータ移行を行う場合、内部の相互結合網及びキャッシュメモリ部１４を介して、２つのディスクアレイ制御ユニット１－１間でデータ移行を行えるため、両方のディスクアレイ制御装置のチャネルＩＦ部を介してデータを移行する必要が無くなり、上記データ移行を実行しているときの性能低下を抑えることが可能となる。したがって、ユーザの通常業務の効率を低下させることがなくなる。

【０１１８】実施例２及び実施例３の構成のディスクアレイ制御装置においても、本実施例を実施する上で問題はなく、本実施例と同様の効果が得られる。

【０１１９】次に実施例１，２，３の実装例について述べる。図１４は、実施例１の図７に示すディスクアレイ制御ユニット１－１を筐体２０１に搭載した一例を示している。

【０１２０】図７に示すチャネルＩＦ部１１はチャネルＩＦパッケージ（ＰＫ）３１１上に実装され、ディスクＩＦ部１２はディスクＩＦＰＫ３１２上に実装され、ＳＭ－ＳＷ１１０およびＣＭ－ＳＷ１１１はスイッチＰＫ３２０上に実装され、共有メモリ部１３及びキャッシュメモリ部１４はメモリＰＫ３３０上に実装されている。また、バックプレーン３４０上にはＳＭアクセスパス１３５、１３６、及びＣＭアクセスパス１３７、１３８が配線されており、バックプレーン３４０に上記各ＰＫを挿す形態となっている。

【０１２１】スイッチＰＫ３２０には、筐体間ＳＭパス１４１用のケーブルと筐体間ＣＭパス１４２用のケーブルが接続され、それぞれのケーブルの他端は、筐体２０１の側面にあるコネクタ２２１、２２２にそれぞれ接続される。ケーブルの図示は省略してある。３５０は電源ボックスであり、上記のＰＫに電力を供給する。このようにディスクアレイ制御ユニットは１つの筐体としてディスクアレイ制御装置の機能を備えている。

【０１２２】図１５は、図１４に示す筐体２０１を２台接続する場合の一例を示している。

【０１２３】スイッチボックス２１０には、図８に示す筐体間ＳＭ－ＳＷ１２１、及び筐体間ＣＭ－ＳＷ１２２が搭載されている。筐体間ＳＭ－ＳＷ１２１に繋がる筐体間ＳＭパス１４１はコネクタ２２１に、筐体間ＣＭ－ＳＷ１２２に繋がる筐体間ＣＭパス１４２はコネクタ２２２に接続されている。

【０１２４】２台の筐体２０１、２０２を接続する場合は、筐体２０１の筐体間ＳＭパス用コネクタ２２１とスイッチボックス２１０のコネクタ２２１をケーブル２３１で、筐体２０１の筐体間ＣＭパス用コネクタ２２２とスイッチボックス２１０のコネクタ２２２をケーブル２３２で繋ぐ。同様に、筐体２０２のコネクタ２２１とスイッチボックス２１０のコネクタ２２１をケーブル２３１で、筐体２０２のコネクタ２２２とスイッチボックス２１０のコネクタ２２２をケーブル２３２で繋ぐ。

【０１２５】上記のようにすることで、１つの筐体でサポートできないホストコンピュータへの接続チャネル数、あるいは１つの筐体でサポートできない記憶容量をサポートすることが可能となる。

【０１２６】ここで、筐体２０１を基本の筐体、筐体２０２を拡張用の筐体とし、スイッチボックス２１０を筐体２０２の中に搭載することも出来る。こうすることにより、基本の筐体２０１の製造コストを上げることなく、スイッチボックス２１０の設置スペースを削除することができる。

【０１２７】実施例２、または実施例３のディスクアレイ制御装置に本実施例を適用することは、何の問題もない。

【０１２８】図１８に筐体間ＳＭ－ＳＷ１２１および筐体間ＣＭ－ＳＷ１２２で３台のディスクアレイ制御ユニットを接続する場合の接続形態を示す。このとき各ＳＷは前述したように接続台数が増えているので２台のディスクアレイ制御ユニットを接続する場合より容量は大きいものが必要とされる。図示のごとく各ディスクアレイ制御ユニット１－１－１～１－１－３は筐体間ＳＭパス１４１および筐体間ＣＭパス１４２によって、筐体間ＳＭ－ＳＷ１２１および筐体間ＣＭ－ＳＷ１２２にそれぞれ接続され、全体として１つのディスクアレイ制御装置として機能する。

【０１２９】図１９にその実装例を示す。ここではスイッチボックス２１０は別筐体として実装されている。これに筐体間ＳＭパス用ケーブル２３１と筐体間ＣＭパス用ケーブル２３２を介してそれぞれコネクタ２２１、コネクタ２２２により筐体２０１，２０２，２０３がそれぞれ接続される。スイッチボックス２１０にディスクアレイ制御ユニットを４台以上接続する容量とコネクタを

用意しておけば後からの増設は容易である。

【０１３０】図２１はスイッチボックス２１０を通るデータのフォーマットを示す。データはパケットの形態を取り、宛先アドレス４０１、コマンド部４０２、データ部４０３からなる。アドレスは共有メモリ、キャッシュメモリ上のアドレスである。

【０１３１】図２２はスイッチボックス２１０内に設けられたスイッチングのためのスイッチ切り替えテーブル４１０を示す。ここには宛先アドレスとそのアドレスを含むディスクアレイ制御ユニットの番号の対応が記憶されている。スイッチボックス２１０はパケット４００のアドレス４０１からこのスイッチ切り替えテーブルを参照し、スイッチの切り替え先を求め、スイッチの切り替え制御を行なう。

【０１３２】ディスクアレイ制御ユニットを増設する場合は次の手順による。スイッチボックス２１０にディスクアレイ制御ユニットを増設するコネクタに余分があれば、そのコネクタにケーブル２３１、ケーブル２３２を接続する。余分がなければスイッチボックスを多段に接続した上でそのコネクタにケーブル２３１、ケーブル２３２を接続する。それと共に、スイッチボックス２１０内のスイッチ切り替えテーブル４１０のアドレスとポートＮｏ．を増設したディスクアレイ制御ユニット分だけ書き加える。前記のアドレスを予め書き込んでおき、増設した場合有効のフラグを立てるやり方もある。

【０１３３】図２０は他の接続例を示すものである。図示のように３つのディスクアレイ制御ユニットが直列に接続されている。このときＳＭ－ＳＷ１１０、ＣＭ－ＳＷ１１１は入力された情報をそのまま他のディスクアレイ制御ユニットに転送するブリッジ機能を持っている。図ではＳＭ－ＳＷ１１０とＣＭ－ＳＷ１１１を有しているが、その代わりにバス構造としても良い。そして、複数のディスクアレイ制御ユニットをバス同士接続した共通バスで結合させても良い。

【０１３４】次に更に他の実装例を示す。図１６に示すように、図１４に示す筐体２０１に搭載したディスクアレイ制御ユニット１－１のパッケージ（ＰＫ）の枚数を減らし、最小構成のディスクアレイ制御ユニット１－１を搭載した筐体２０５とする。

【０１３５】図１７に示すように、１つの筐体２０６内に２個以上の筐体２０５とスイッチボックス２１０を搭載し、筐体２０５間を実施例６で示した方法と同様の方法で、スイッチボックス２１０を介して接続することにより、中規模から大規模な構成のディスクアレイ制御装置を構成することが可能となる。

【０１３６】なお、筐体２０５はいままで述べてきたディスクアレイ制御ユニットの機能を持っているものであれば良く、筐体の形をとらずにモジュールと呼ばれる形態をとっても良い。また、図１７の場合ディスクアレイ制御ユニット毎に電源ボックスを持たせるか、共通の電源ボックスから給電するかは実装上の事柄として適宜決められるものである。

【０１３７】

【発明の効果】本発明によれば、複数台のディスクアレイ制御装置を１つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置間でのデータ移行による性能低下を抑えるディスクアレイシステムを提供することが可能となる。

【図面の簡単な説明】

【図１】本発明によるディスクアレイ制御装置の構成を示す図。

【図２】従来のディスクアレイ制御装置の構成を示す図。

【図３】従来のディスクアレイ制御装置の他の構成を示す図。

【図４】従来のディスクアレイ制御装置の他の構成を示す図。

【図５】本発明によるディスクアレイ制御装置の他の構成を示す図。

【図６】本発明によるディスクアレイ制御装置の他の構成を示す図。

【図７】図１に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図８】図７に示すディスクアレイ制御ユニットを複数台接続する構成を示す図。

【図９】図５に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図１０】図９に示すディスクアレイ制御ユニットを複数台接続する構成を示す図。

【図１１】図６に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図１２】図１１に示すディスクアレイ制御ユニットを複数台接続する構成を示す図。

【図１３】図７に示すディスクアレイ制御ユニットを複数台接続する他の構成を示す図。

【図１４】本発明によるディスクアレイ制御ユニットの筐体への搭載例を示す図。

【図１５】本発明によるディスクアレイ制御ユニットを搭載した筐体を複数台接続する構成を示す図。

【図１６】本発明によるディスクアレイ制御ユニットの筐体への搭載の他の例を示す図。

【図１７】本発明によるディスクアレイ制御ユニットを複数台、１つの筐体へ搭載する例を示す図。

【図１８】筐体間スイッチにより３台のディスクアレイ制御システムを接続する配線構造を示す図。

【図１９】図１８の配線構造の一実装例を示す図。

【図２０】本発明によるディスクアレイ制御ユニットを３台以上接続するための他の実施例を示す図。

【図２１】スイッチボックスに与えられる情報のデータフォーマットの一例を示す図。

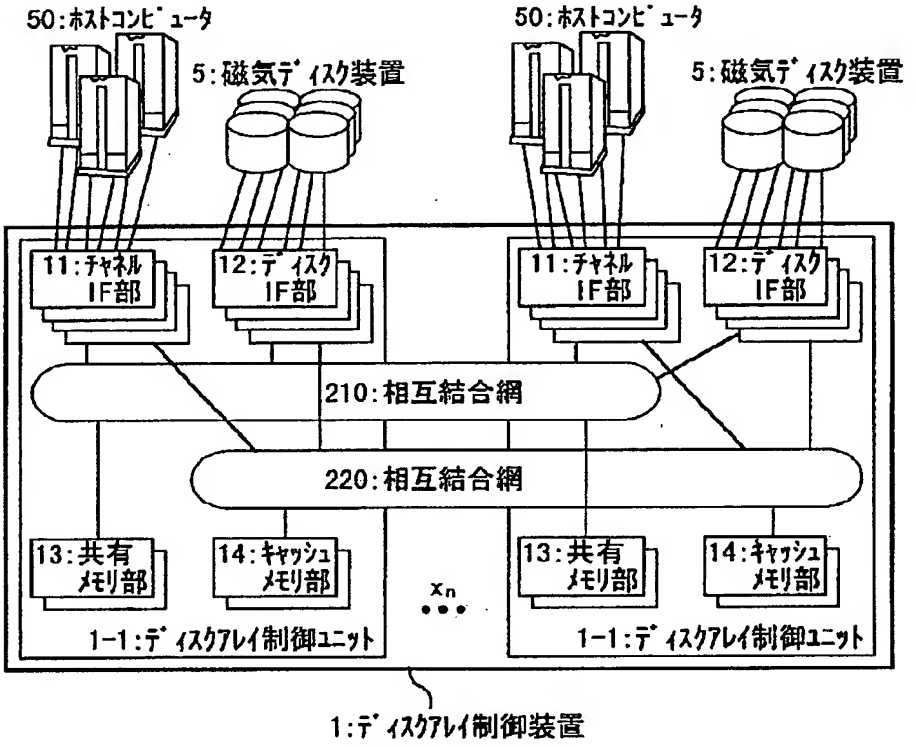
【図２２】スイッチボックス内に設けられたスイッチ切り替えテーブルの一例を示す図。

【符号の説明】

１：ディスクアレイ制御装置、１－１…ディスクアレイ制御ユニット、５…磁気ディスク装置、１１…チャネルＩＦ部、１２…ディスクＩＦ部、１３…共有メモリ部、１４…キャッシュメモリ部、２１０、２２０…相互結合網、５０…ホストコンピュータ

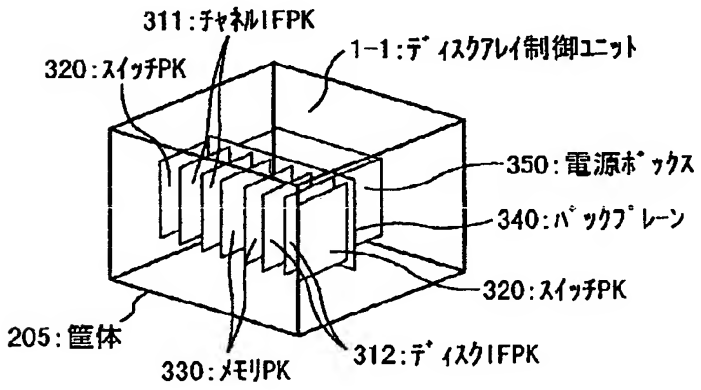
【図１】

図　1



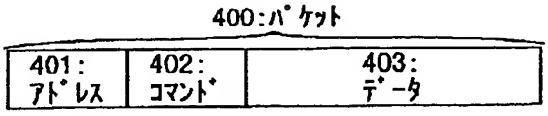
【図１６】

図　16



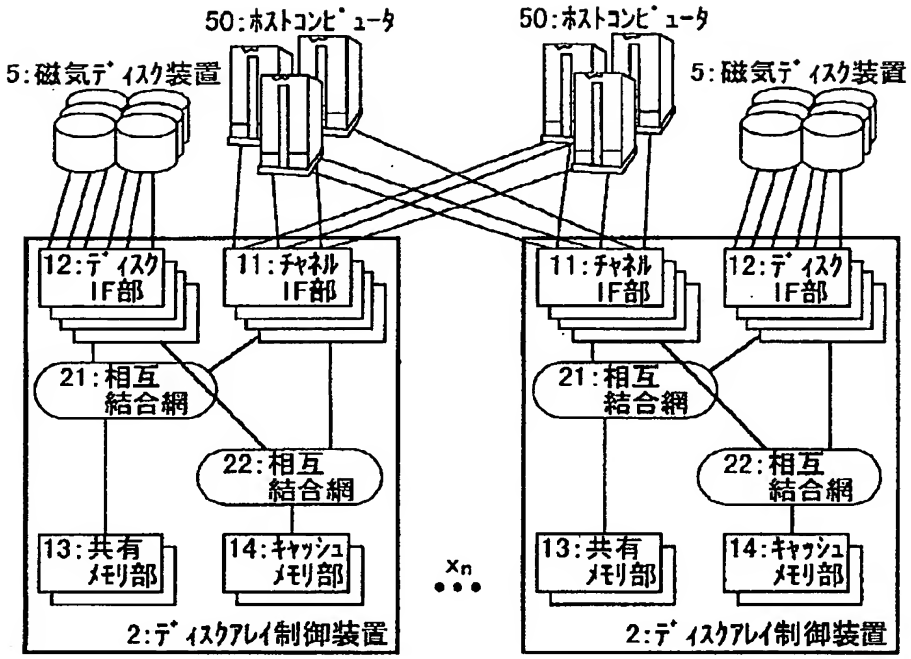
【図２１】

図　21



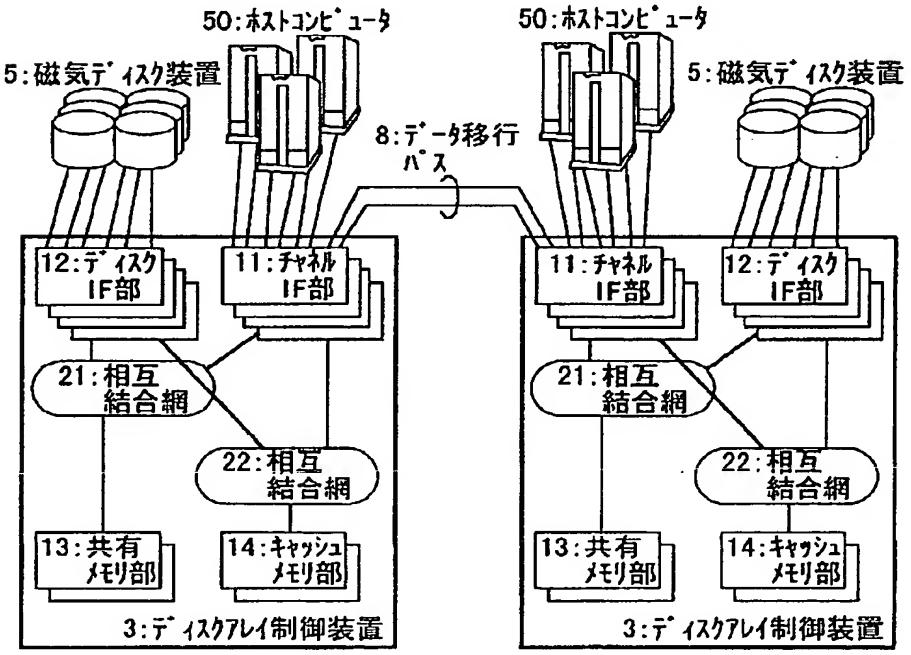
【図２】

図　2


50:ホストコンピュータ
50:ホストコンピュータ
5:磁気ディスク装置
5:磁気ディスク装置
12:ディスク IF部
11:チャネル IF部
11:チャネル IF部
12:ディスク IF部
21:相互結合網
21:相互結合網
22:相互結合網
22:相互結合網
13:共有メモリ部
14:キャッシュメモリ部
13:共有メモリ部
14:キャッシュメモリ部
Xn
2:ディスクアレイ制御装置
2:ディスクアレイ制御装置

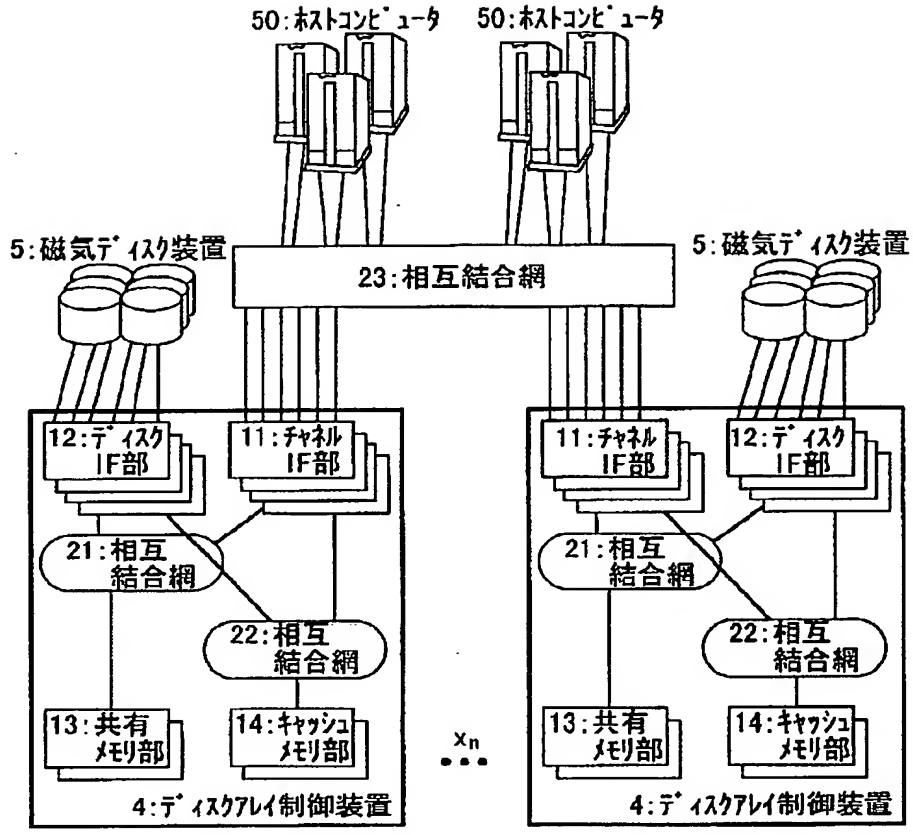
【図３】

図　3


50:ホストコンピュータ
50:ホストコンピュータ
5:磁気ディスク装置
5:磁気ディスク装置
8:データ移行パス
12:ディスク IF部
11:チャネル IF部
11:チャネル IF部
12:ディスク IF部
21:相互結合網
21:相互結合網
22:相互結合網
22:相互結合網
13:共有メモリ部
14:キャッシュメモリ部
13:共有メモリ部
14:キャッシュメモリ部
3:ディスクアレイ制御装置
3:ディスクアレイ制御装置

【図４】

図 4



【図２２】

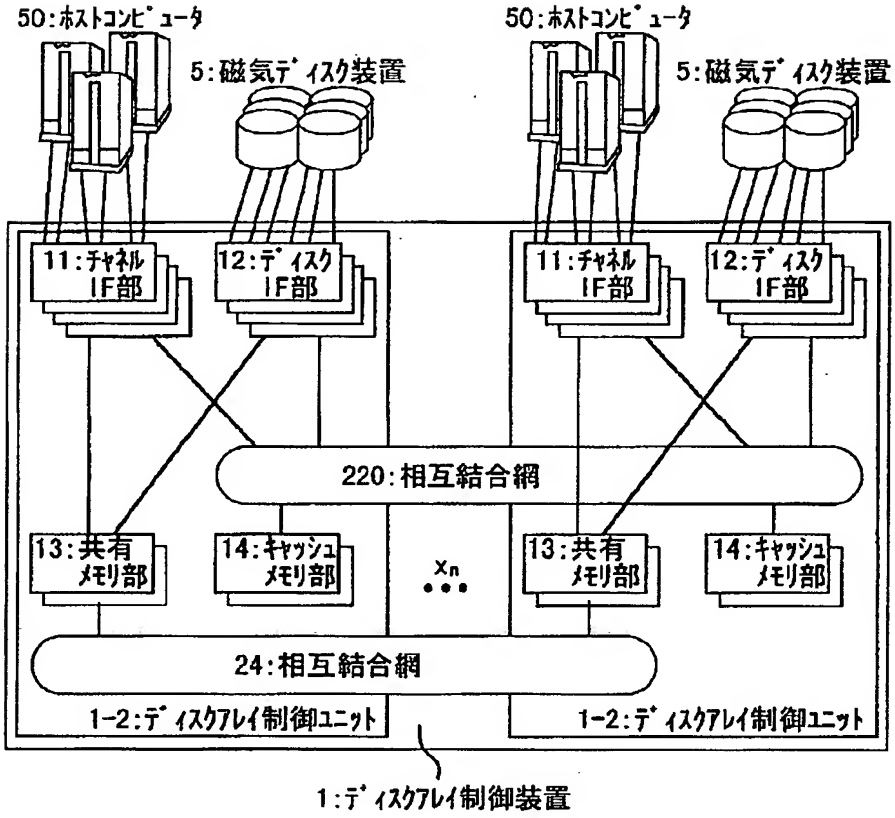
図 22

410:スイッチ切替テーブル

アドレス	ポートNo
0×0…000	0
0×0…001	0
⋮	⋮
0×0…0FF	0
0×0…100	1
⋮	⋮
0×0…200	2
⋮	⋮

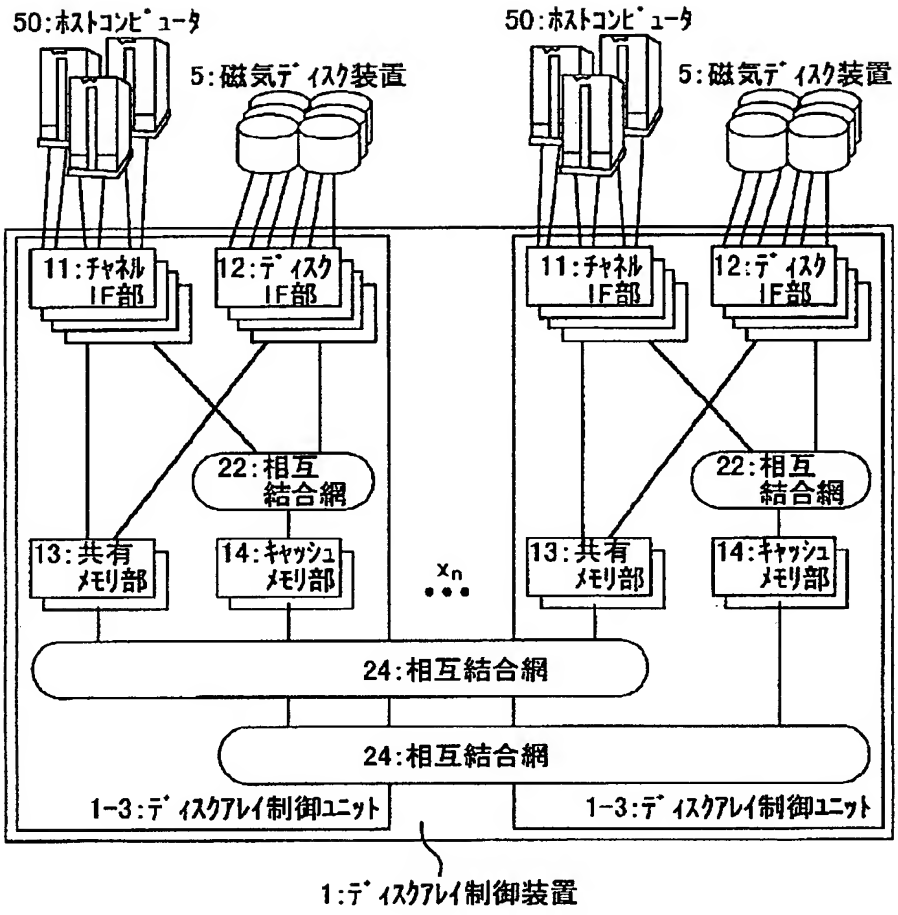
【図５】

図 5


50:ホストコンピュータ
5:磁気ディスク装置
50:ホストコンピュータ
5:磁気ディスク装置
11:チャネル IF部
12:ディスク IF部
11:チャネル IF部
12:ディスク IF部
220:相互結合網
13:共有 メモリ部
14:キャッシュ メモリ部
xn
13:共有 メモリ部
14:キャッシュ メモリ部
24:相互結合網
1-2:ディスクアレイ制御ユニット
1-2:ディスクアレイ制御ユニット
1:ディスクアレイ制御装置

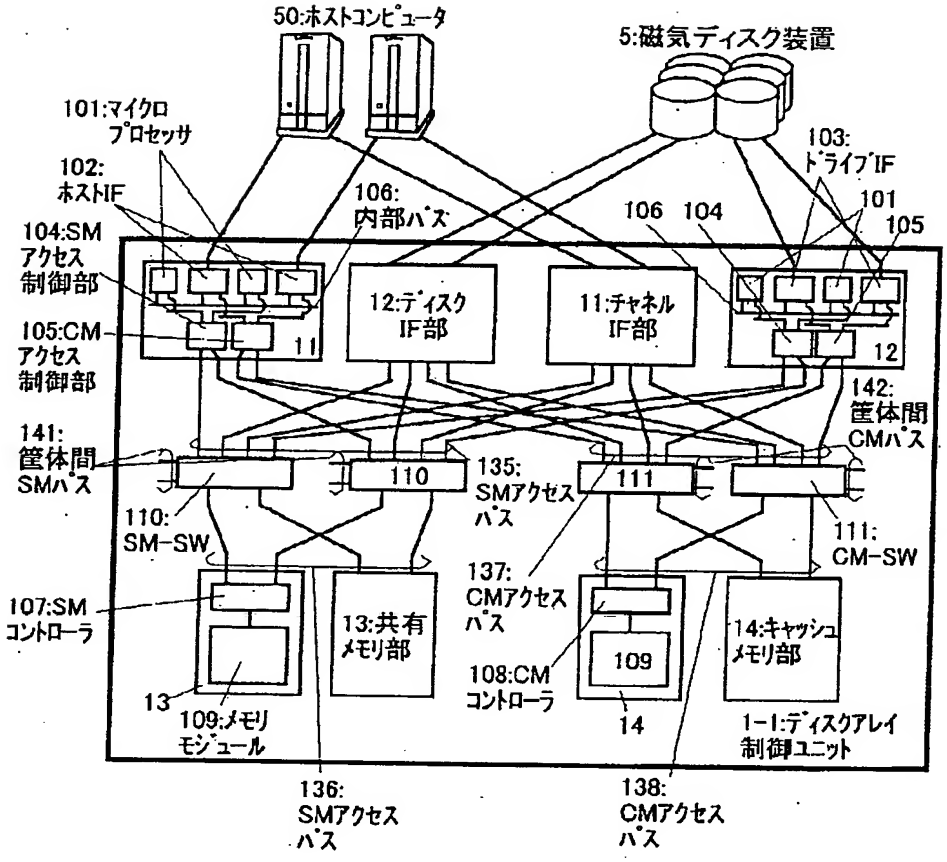
【図６】

図 6


50:ホストコンピュータ
5:磁気ディスク装置
50:ホストコンピュータ
5:磁気ディスク装置
11:チャネル IF部
12:ディスク IF部
11:チャネル IF部
12:ディスク IF部
22:相互結合網
22:相互結合網
13:共有メモリ部
14:キャッシュメモリ部
xn
13:共有メモリ部
14:キャッシュメモリ部
24:相互結合網
24:相互結合網
1-3:ディスクアレイ制御ユニット
1-3:ディスクアレイ制御ユニット
1:ディスクアレイ制御装置

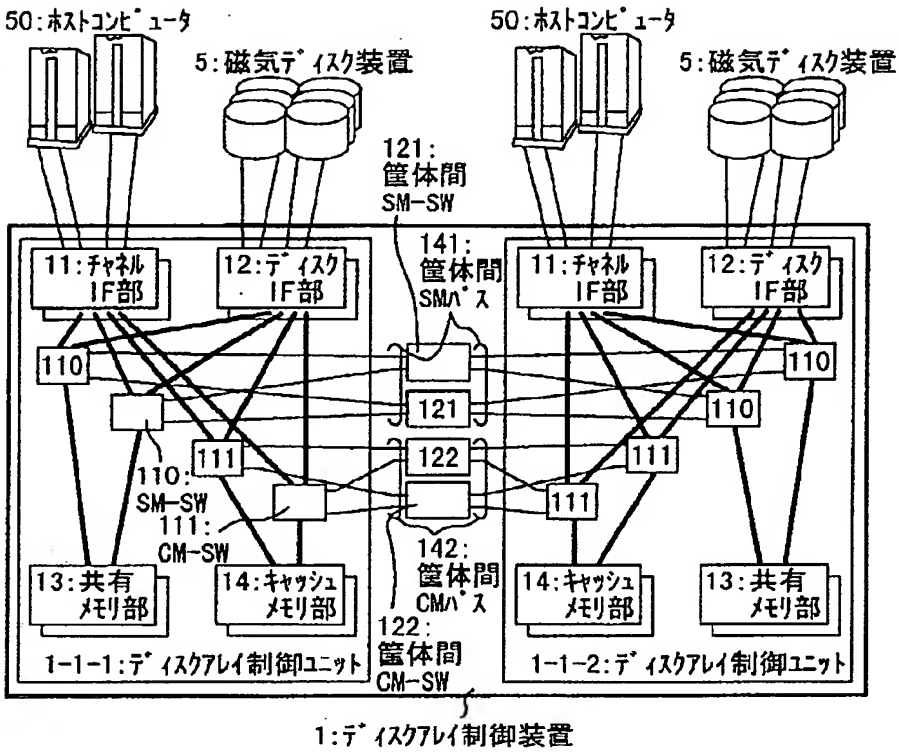
【図７】

図7


50:ホストコンピュータ
5:磁気ディスク装置
101:マイクロ
プロセッサ
102:
ホストIF
106:
内部パス
103:
ドライブIF
101
105
106
104
104:SM
アクセス
制御部
105:CM
アクセス
制御部
11
12:ディスク
IF部
11:チャネル
IF部
12
142:
筐体間
CMパス
141:
筐体間
SMパス
110
135:
SMアクセス
パス
111
110:
SM-SW
111:
CM-SW
137:
CMアクセス
パス
107:SM
コントローラ
13:共有
メモリ部
14:キャッシュ
メモリ部
109
108:CM
コントローラ
13
109:メモリ
モジュール
14
1-1:ディスクアレイ
制御ユニット
136:
SMアクセス
パス
138:
CMアクセス
パス

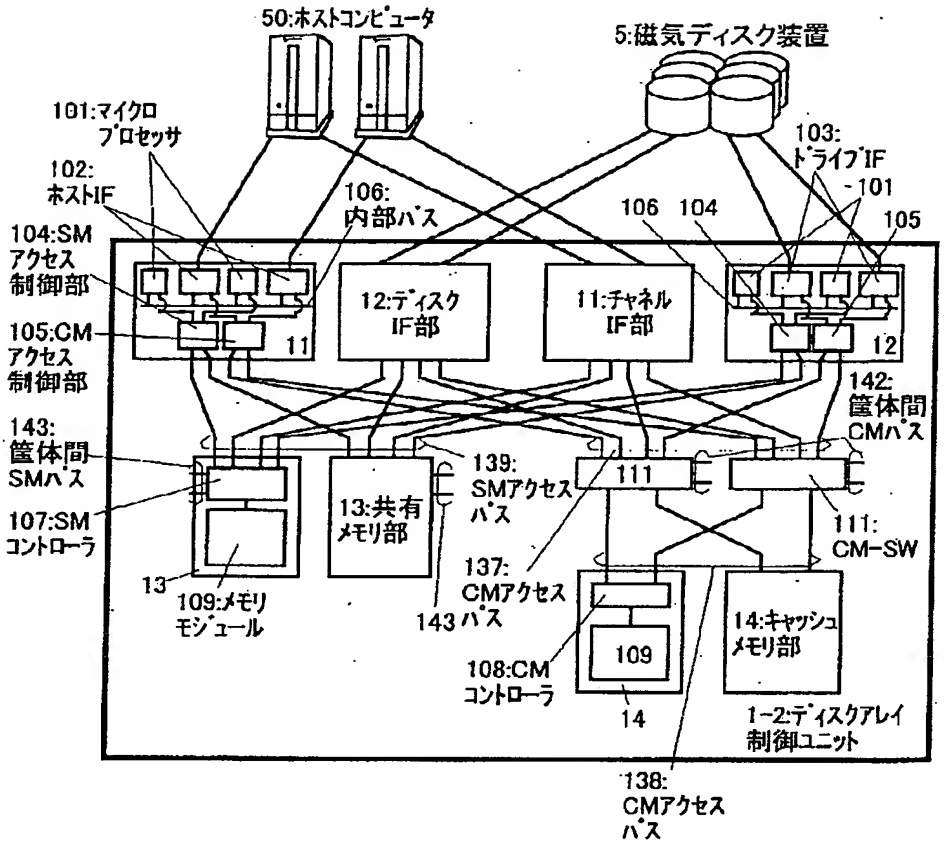
【図８】

図　8


50:ホストコンピュータ
5:磁気ディスク装置
50:ホストコンピュータ
5:磁気ディスク装置
121:
筐体間
SM-SW
141:
筐体間
SMパス
11:チャネル
IF部
12:ディスク
IF部
11:チャネル
IF部
12:ディスク
IF部
110
110
121
110
111
122
111
110:
SM-SW
111:
CM-SW
111
142:
筐体間
CMパス
13:共有
メモリ部
14:キャッシュ
メモリ部
14:キャッシュ
メモリ部
13:共有
メモリ部
122:
筐体間
CM-SW
1-1-1:ディスクアレイ制御ユニット
1-1-2:ディスクアレイ制御ユニット
1:ディスクアレイ制御装置

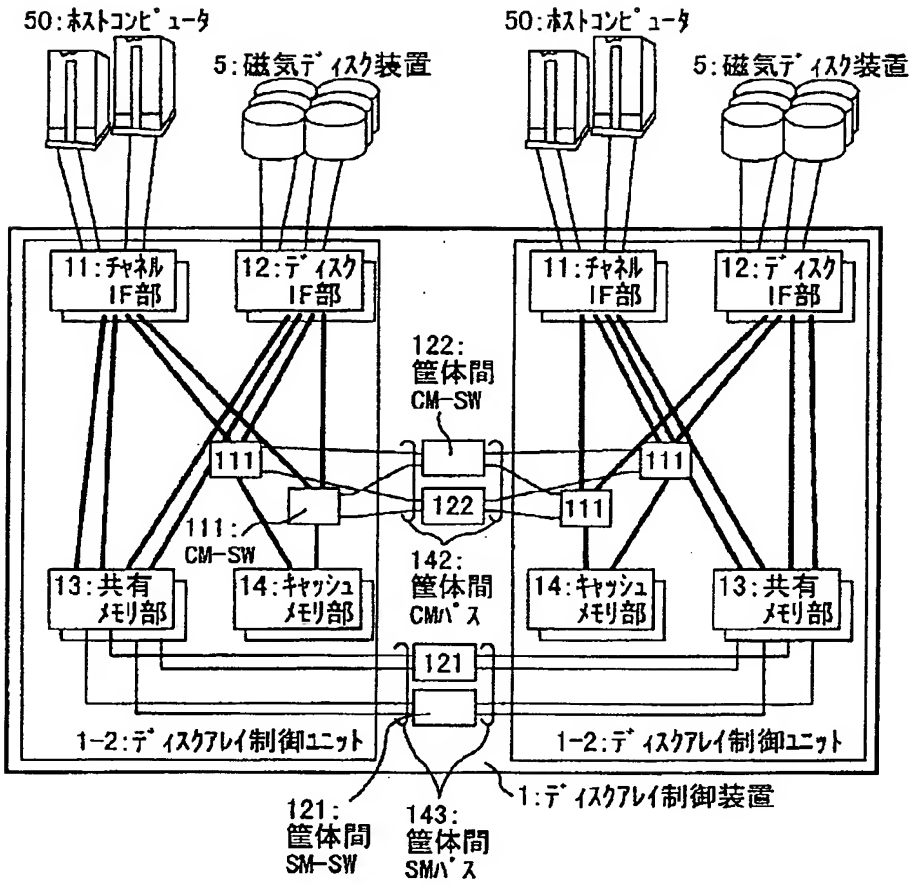
【図９】

図9


50:ホストコンピュータ
5:磁気ディスク装置
101:マイクロプロセッサ
102:ホストIF
106:内部パス
103:ドライブIF
101
106
104
105
104:SMアクセス制御部
105:CMアクセス制御部
11
12:ディスクIF部
11:チャネルIF部
12
142:筐体間CMパス
143:筐体間SMパス
139:SMアクセスパス
111
107:SMコントローラ
13:共有メモリ部
111:CM-SW
13
109:メモリモジュール
137:CMアクセスパス
143
108:CMコントローラ
109
14:キャッシュメモリ部
14
1-2:ディスクアレイ制御ユニット
138:CMアクセスパス

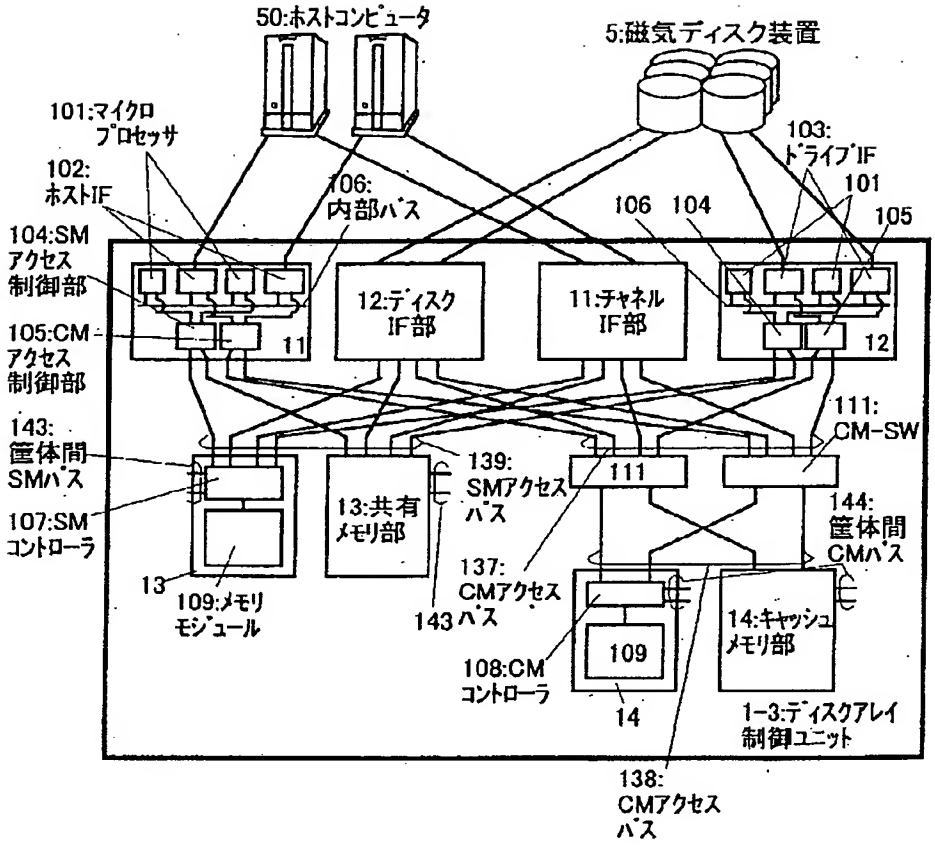
【図１０】

図　10


50:ホストコンピュータ
5:磁気ディスク装置
50:ホストコンピュータ
5:磁気ディスク装置
11:チャネル IF部
12:ディスク IF部
11:チャネル IF部
12:ディスク IF部
122: 筐体間 CM-SW
111
111
122
111
111: CM-SW
142: 筐体間 CMパス
13:共有 メモリ部
14:キャッシュ メモリ部
14:キャッシュ メモリ部
13:共有 メモリ部
121
1-2:ディスクアレイ制御ユニット
1-2:ディスクアレイ制御ユニット
1:ディスクアレイ制御装置
121: 筐体間 SM-SW
143: 筐体間 SMパス

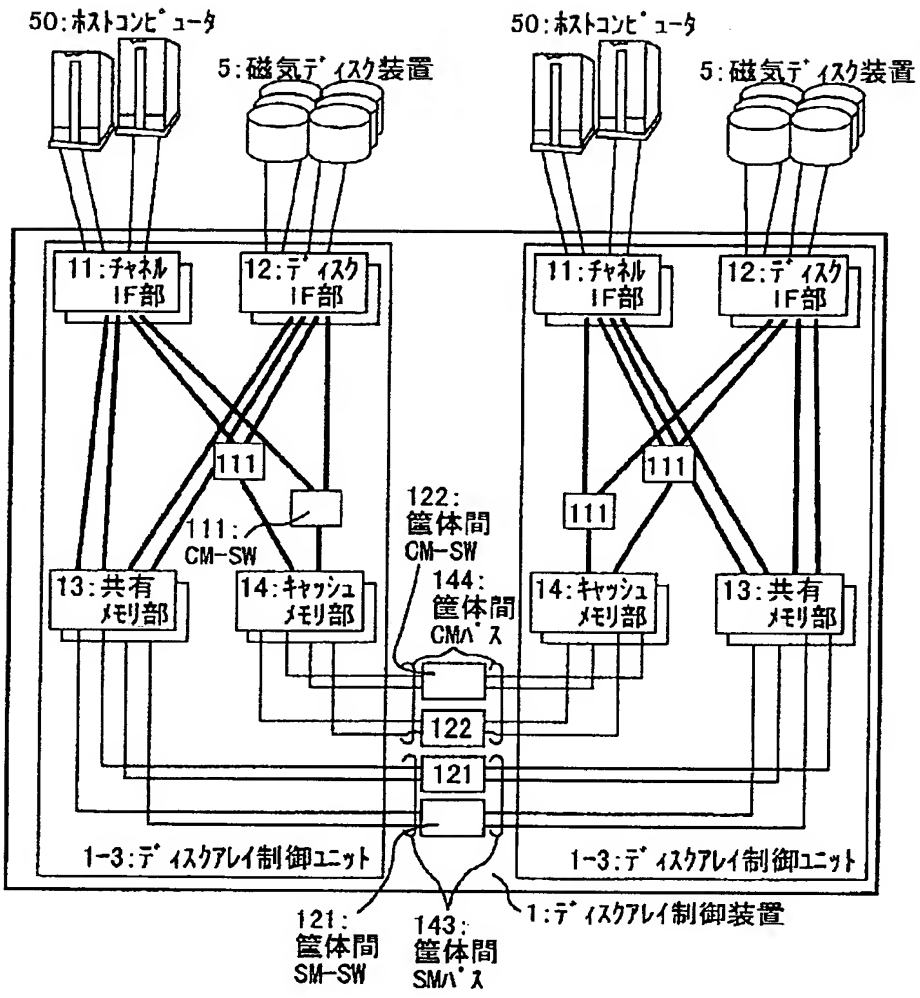
【図１１】

図11


50:ホストコンピュータ
5:磁気ディスク装置
101:マイクロプロセッサ
102:ホストIF
106:内部バス
103:ドライブIF
101
106
104
105
104:SMアクセス制御部
105:CMアクセス制御部
11
12:ディスクIF部
11:チャネルIF部
12
111:CM-SW
143:筐体間SMパス
107:SMコントローラ
139:SMアクセスパス
111
144:筐体間CMパス
13:共有メモリ部
137:CMアクセスパス
13
109:メモリモジュール
143
109
14:キャッシュメモリ部
108:CMコントローラ
14
1-3:ディスクアレイ制御ユニット
138:CMアクセスパス

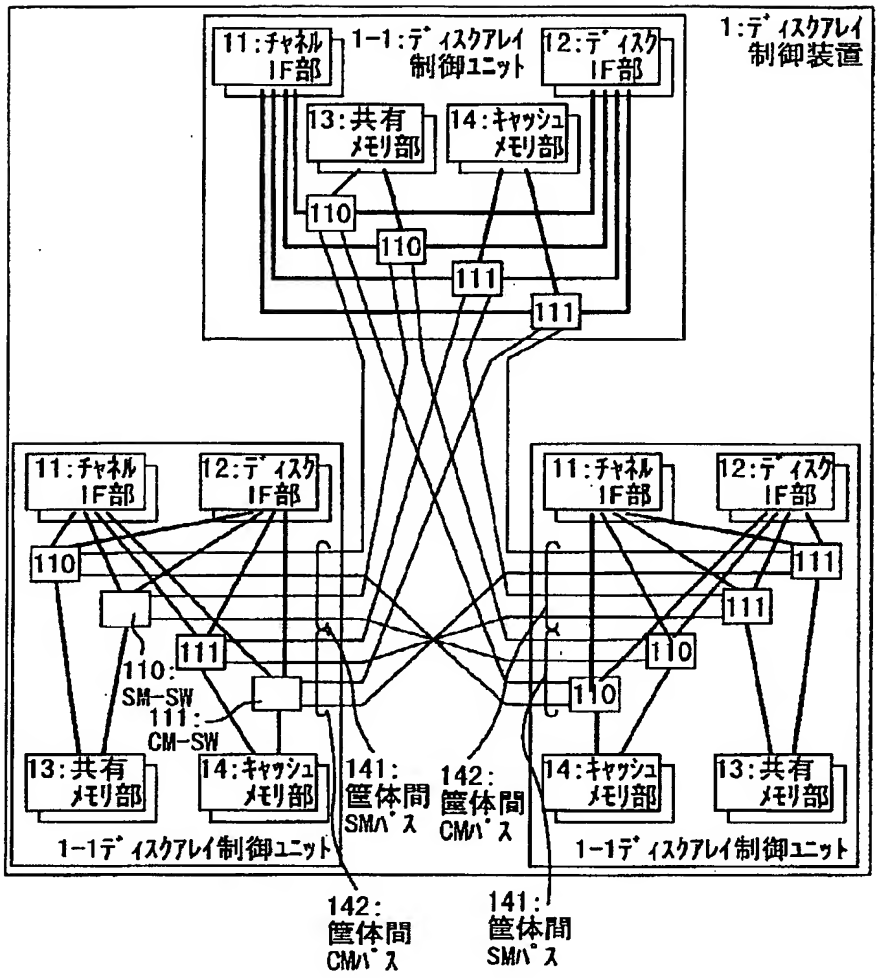
【図１２】

図 12


50:ホストコンピュータ
5:磁気ディスク装置
50:ホストコンピュータ
5:磁気ディスク装置
11:チャネル IF部
12:ディスク IF部
11:チャネル IF部
12:ディスク IF部
111
111
111:
CM-SW
122:
筐体間
CM-SW
144:
筐体間
CMパス
111
14:キャッシュ メモリ部
13:共有 メモリ部
14:キャッシュ メモリ部
13:共有 メモリ部
122
121
1-3:ディスクアレイ制御ユニット
1-3:ディスクアレイ制御ユニット
1:ディスクアレイ制御装置
121:
筐体間
SM-SW
143:
筐体間
SMパス

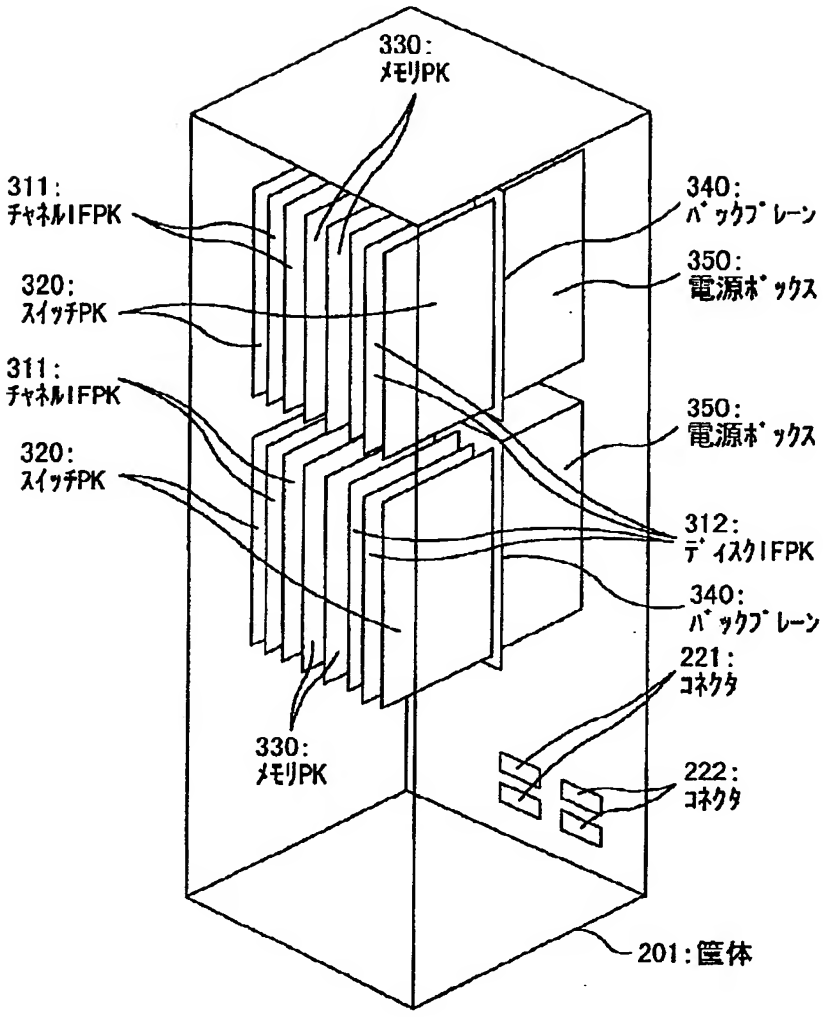
【図１３】

図 13


11:チャネル IF部
1-1:ディスクアレイ 制御ユニット
12:ディスク IF部
1:ディスクアレイ 制御装置
13:共有 メモリ部
14:キャッシュ メモリ部
110
111
110:
SM-SW
111:
CM-SW
141: 筐体間 SMパス
142: 筐体間 CMパス
1-1ディスクアレイ制御ユニット

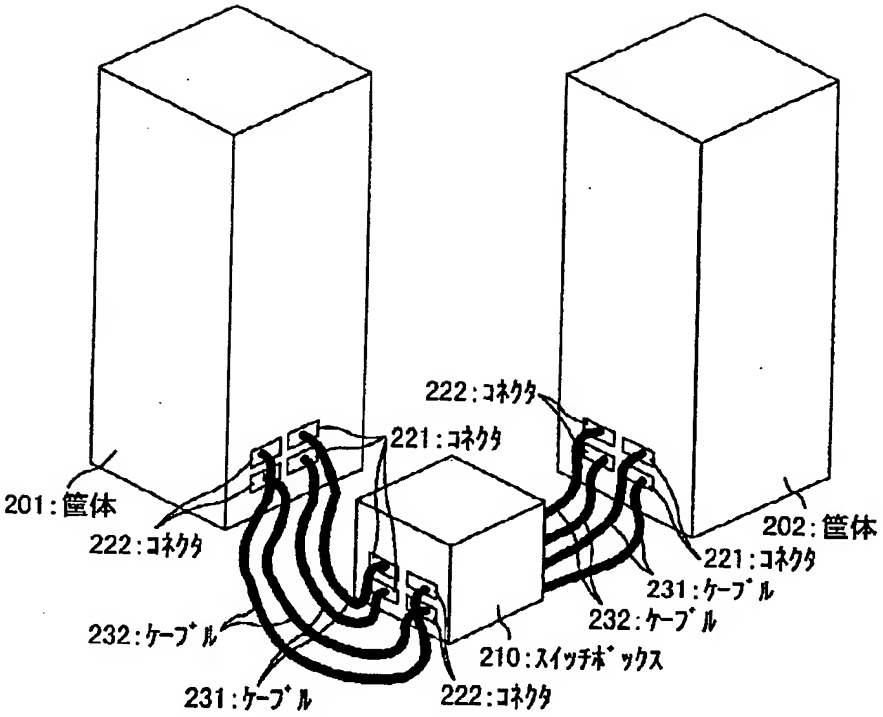
【図１４】

図　14


330:
メモリPK
311:
チャネルIFPK
320:
スイッチPK
311:
チャネルIFPK
320:
スイッチPK
330:
メモリPK
340:
バックプレーン
350:
電源ボックス
350:
電源ボックス
312:
ディスクIFPK
340:
バックプレーン
221:
コネクタ
222:
コネクタ
201:筐体

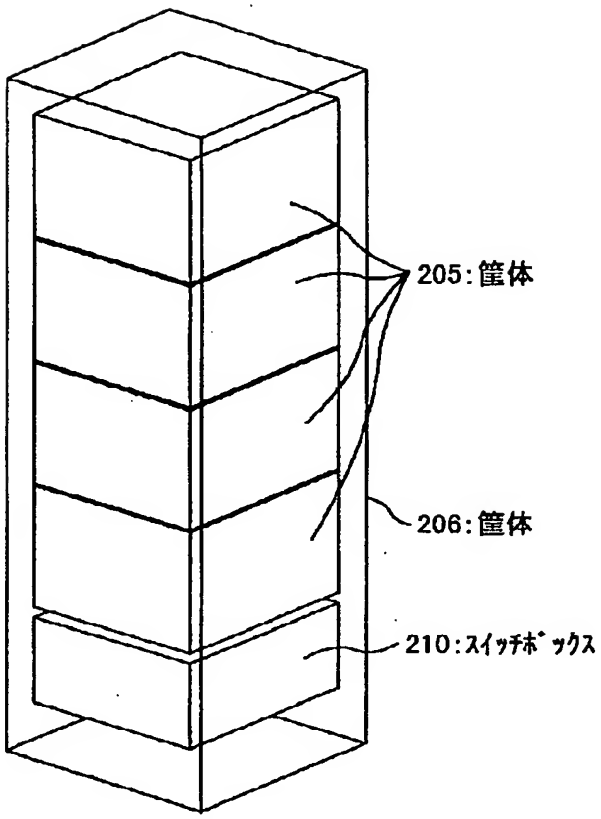
【図１５】

図　15


222:ｺﾈｸﾀ
221:ｺﾈｸﾀ
201:筐体
222:ｺﾈｸﾀ
202:筐体
221:ｺﾈｸﾀ
231:ｹｰﾌﾞﾙ
232:ｹｰﾌﾞﾙ
232:ｹｰﾌﾞﾙ
210:ｽｲｯﾁﾎﾞｯｸｽ
231:ｹｰﾌﾞﾙ
222:ｺﾈｸﾀ

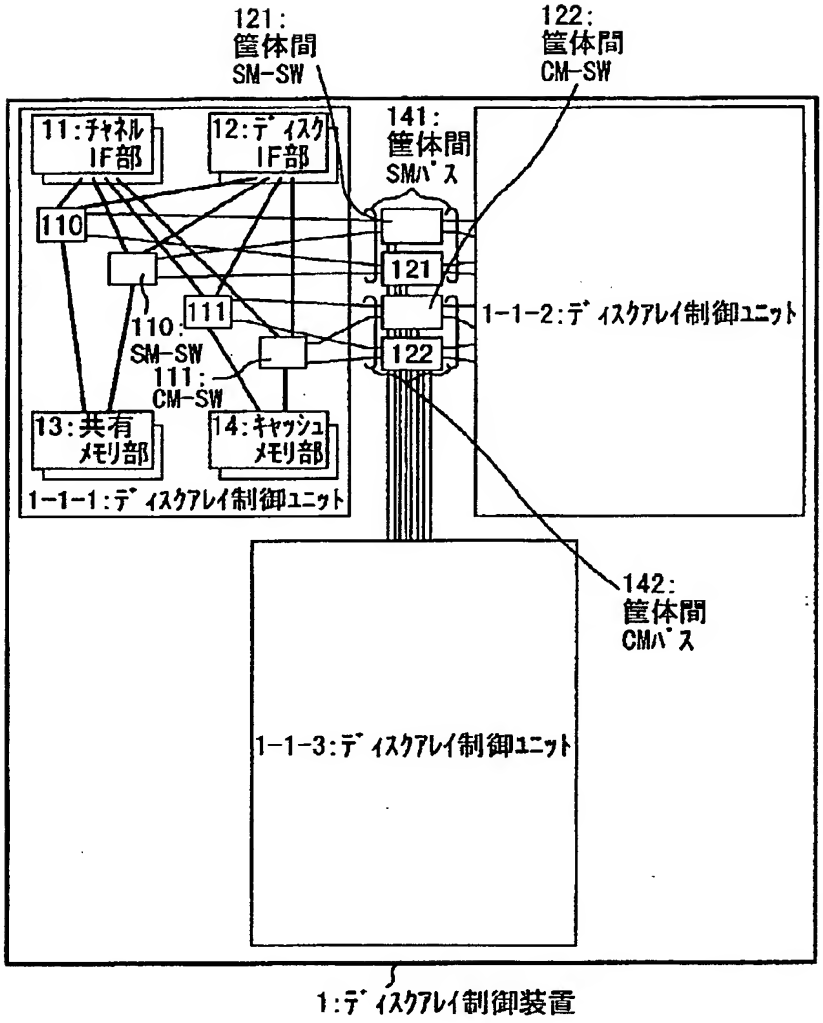
【図１７】

図 17



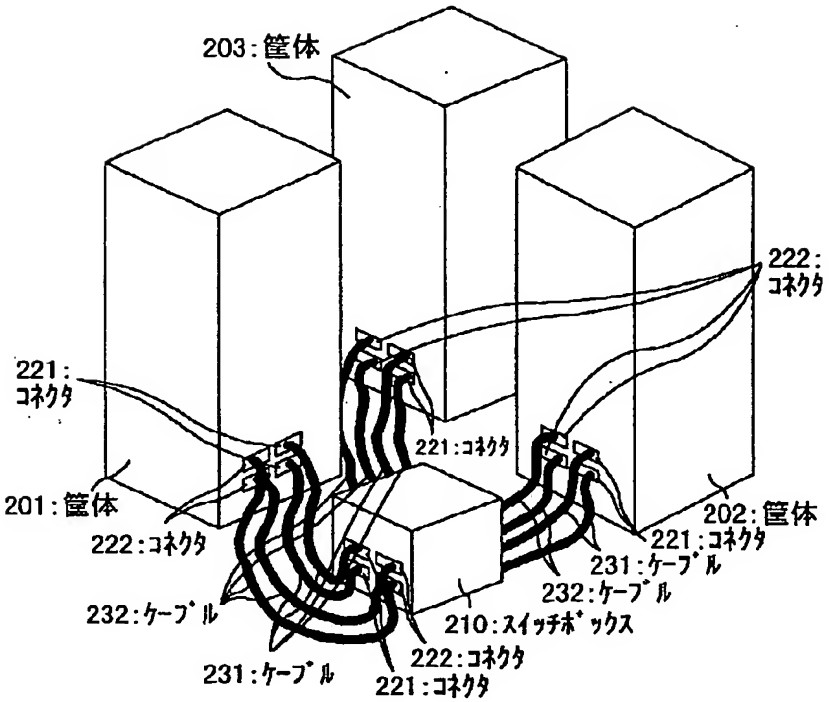
【図１８】

図 18


121:
筐体間
SM-SW
122:
筐体間
CM-SW
11:チャネル
IF部
12:ディスク
IF部
141:
筐体間
SMパス
110
121
111
122
110:
SM-SW
111:
CM-SW
1-1-2:ディスクアレイ制御ユニット
13:共有
メモリ部
14:キャッシュ
メモリ部
1-1-1:ディスクアレイ制御ユニット
142:
筐体間
CMパス
1-1-3:ディスクアレイ制御ユニット
1:ディスクアレイ制御装置

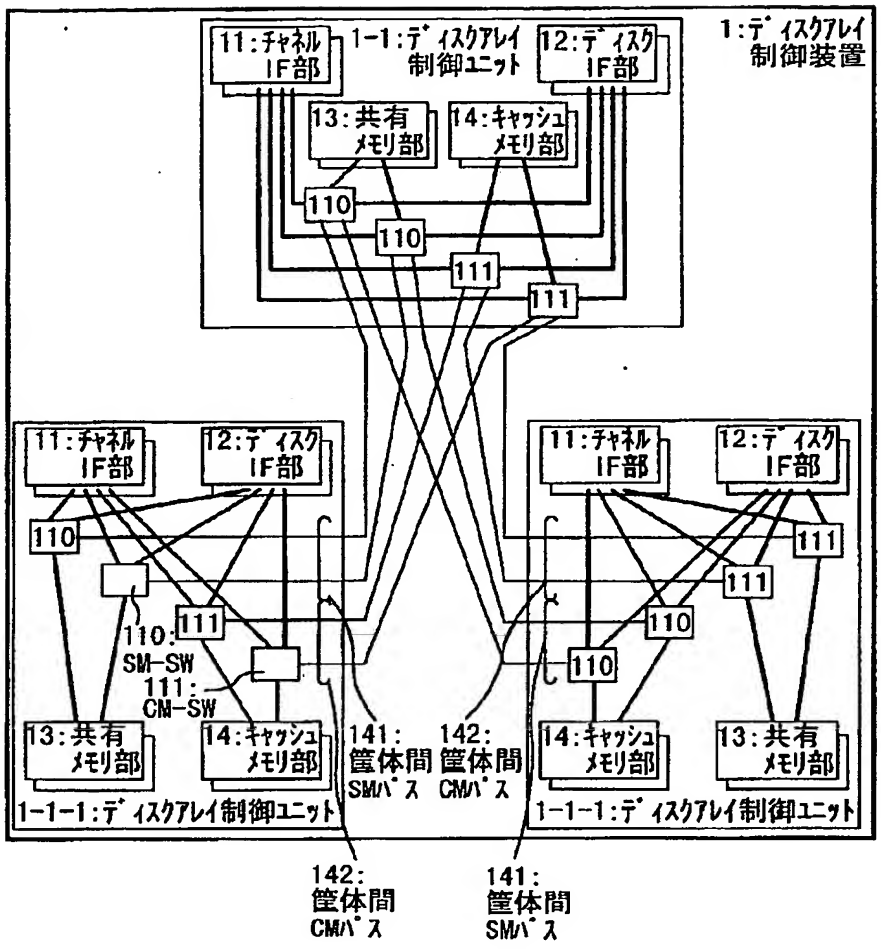
【図１９】

図　19


203:筐体
222:
ｺﾈｸﾀ
221:
ｺﾈｸﾀ
221:ｺﾈｸﾀ
201:筐体
222:ｺﾈｸﾀ
202:筐体
221:ｺﾈｸﾀ
231:ｹｰﾌﾞﾙ
232:ｹｰﾌﾞﾙ
232:ｹｰﾌﾞﾙ
210:ｽｲｯﾁﾎﾞｯｸｽ
222:ｺﾈｸﾀ
231:ｹｰﾌﾞﾙ
221:ｺﾈｸﾀ

【図２０】

図　20



フロントページの続き

(51)Int.Cl.⁷	識別記号	FI	ﾃｰﾏｺｰﾄﾞ(参考)
G06F 12/16	320	G06F 12/16	320L

(72)発明者　藤林　昭
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者　櫻井　亘
神奈川県小田原市国府津2880番地　株式会
社日立製作所ストレージシステム事業部内

Fターム(参考)　5B005 JJ01 MM12 WW02 WW04
5B018 GA04 HA04 MA14
5B065 BA01 CA30 CE22 CE26 CH01
CH13